

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:18:03 ; Search time 166 Seconds
(without alignments)
382.101 Million cell updates/sec

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARLNLKKKTNTVPTVAHRT.....ALPGPWRWASGFWKATMQQ 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	837	97.7	160	8	ADG22790	Adg22790 Hepatitis
2	813.5	94.9	198	3	AAY44545	Aay44545 Hepatitis
3	806	94.0	161	6	ABU61864	Abu61864 HCV core
4	768	89.6	161	6	ABU61866	Abu61866 HCV core
5	763	89.0	161	6	ABU61865	Abu61865 HCV core
6	645.5	75.3	198	3	AAY44558	Aay44558 Hepatitis
7	624	72.8	161	6	ABU61870	Abu61870 HCV core
8	583	68.0	198	3	AAY44554	Aay44554 Hepatitis
9	552	64.4	143	6	ABU61868	Abu61868 HCV core
10	548	63.9	143	6	ABU61872	Abu61872 HCV core
11	547	63.8	143	6	ABU61871	Abu61871 HCV core
12	540	63.0	143	6	ABU61867	Abu61867 HCV core
13	540	63.0	143	6	ABU61873	Abu61873 HCV core
14	540	63.0	143	6	ABU61874	Abu61874 HCV core
15	534	62.3	143	6	ABU61875	Abu61875 HCV core
16	508	59.3	198	3	AAY44557	Aay44557 Hepatitis
17	501	58.5	198	3	AAY44552	Aay44552 Hepatitis
18	487	56.8	154	6	ABU61879	Abu61879 HCV core
19	464	54.1	198	3	AAY44556	Aay44556 Hepatitis
20	441	51.5	198	3	AAY44553	Aay44553 Hepatitis
21	428	49.9	139	6	ABU61869	Abu61869 HCV core
22	426	49.7	99	8	ADR22073	Adr22073 Anti-Hepa
23	425	49.6	99	8	ADR21979	Adr21979 Anti-Hepa
24	424	49.5	154	6	ABU61878	Abu61878 HCV core
25	422	49.2	99	8	ADR21949	Adr21949 Anti-Hepa

26	421	49.1	99	8	ADR22067	Adr22067 Anti-Hepa
27	421	49.1	99	8	ADR21937	Adr21937 Anti-Hepa
28	420	49.0	99	8	ADR21997	Adr21997 Anti-Hepa
29	419	48.9	99	8	ADR21978	Adr21978 Anti-Hepa
30	417	48.7	99	8	ADR21944	Adr21944 Anti-Hepa
31	417	48.7	139	6	ABU61880	Abu61880 HCV core
32	416	48.5	99	8	ADR21959	Adr21959 Anti-Hepa
33	416	48.5	99	8	ADR21948	Adr21948 Anti-Hepa
34	415	48.4	99	8	ADR22033	Adr22033 Anti-Hepa
35	415	48.4	99	8	ADR22072	Adr22072 Anti-Hepa
36	415	48.4	99	8	ADR22065	Adr22065 Anti-Hepa
37	415	48.4	99	8	ADR21947	Adr21947 Anti-Hepa
38	415	48.4	99	8	ADR22029	Adr22029 Anti-Hepa
39	415	48.4	99	8	ADR21993	Adr21993 Anti-Hepa
40	415	48.4	99	8	ADR21950	Adr21950 Anti-Hepa
41	414	48.3	99	8	ADR22070	Adr22070 Anti-Hepa
42	414	48.3	99	8	ADR21945	Adr21945 Anti-Hepa
43	414	48.3	99	8	ADR22038	Adr22038 Anti-Hepa
44	414	48.3	99	8	ADR21960	Adr21960 Anti-Hepa
45	413	48.2	99	8	ADR21986	Adr21986 Anti-Hepa

ALIGNMENTS

RESULT 1

ADG22790
ID ADG22790 standard; protein; 160 AA.

AC ADG22790;

DT 26-FEB-2004 (first entry)

DE Hepatitis C virus F protein.

KW virucide; antiviral RNA silencing pathway modulator; gene inactivation;

RNA silencing; recombinant construct; antiviral RNA silencing pathway;

viral infection; hepatitis C virus; F protein; RNA silencing suppressor.

OS Hepatitis C virus.

FN US2003219407-A1.

PD 27-NOV-2003.

PF 15-MAY-2002; 2002US-00150283.

PR 15-MAY-2002; 2002US-00150283.

(REGC) UNIV CALIFORNIA.

Ding S, Li H, Li W;

WPI; 2004-021927/02.

N-PSDB; ADG22791.

New recombinant DNA construct comprising a polynucleotide sequence of a virus which activates RNA silencing, and a polynucleotide sequence of a gene, useful for inactivating a gene in a cell.

Disclosure; SEQ ID NO 1; 20pp; English.

The invention describes a recombinant DNA construct for inactivation of a gene in a cell. The construct comprises: a polynucleotide sequence of a virus which activates RNA silencing; and a polynucleotide sequence of a gene. The recombinant constructs are useful in inactivating a gene in a cell or in an animal. Modulators of antiviral RNA silencing pathway are useful for treating or preventing viral infection. This is the amino acid sequence of hepatitis C virus F protein, a RNA silencing suppressor.

SQ Sequence 160 AA;

Query Match 97.7%; Score 837; DB 8; Length 160;

Best Local Similarity 100.0%; Pred. No. 5.1e-77;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRERLPSGRNLEVD 60
DB 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRERLPSGRNLEVD 60
QY 61 VLSLPRVLGPRAGPGLSPGTLGSPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWR 120
DB 61 VLSLPRVLGPRAGPGLSPGTLGSPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPLRAASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 160
DB 121 SSIPLRAASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 160

RESULT 2
AAI44545
ID AAY44545 standard; protein; 198 AA.
XX AAY44545;
XX AC AAY44545;
XX DT 04-APR-2000 (first entry)
XX DE Hepatitis C virus protein encoded by DNA clone AF011751.
XX KW Hepatitis C virus protein; HCV; hepatitis C; immunogenic; vaccine;
XX KW prevention; diagnosis; therapeutic target; anti-HCV therapy;
XX KW HCV infection; DNA clone AF011751.
XX OS Hepatitis C virus.
XX FH Key Location/Qualifiers
FT Misc-difference 146 /note= "Encoded by CTG"
FT FT Misc-difference 161 /label= unknown
FT FT Misc-difference 184 /note= "Encoded by TGA"
FT FT Misc-difference 184 /label= unknown
FT FT /note= "Encoded by TGA"
XX W09963941-A2.
XX 16-DEC-1999.
XX 09-JUN-1999; 99WO-US012929.
XX 09-JUN-1998; 98US-0088670P.
XX 11-JUN-1998; 98US-0089138P.
XX (BRAN/) BRANCH A D.
XX (WALE/) WALEWSKI J L.
XX (STUM/) STUMP D D.
XX Branch AD, Walewski JL, Stump DD;
XX WPI; 2000-126431/11.
XX N-PSDB; AAZ29897.
XX Novel Hepatitis C virus peptides useful in vaccine compositions, for
XX diagnosing HCV infection and as therapeutic agents.
XX Claim 1; Page 39-41; 50pp; English.
XX The present sequence is a Hepatitis C virus (HCV) protein derived from an
XX infectious HCV isolate AF011751. The novel HCV protein is encoded by a
XX reading frame +1 or +2 relative to the standard HCV open reading frame
XX hence not derived from the standard HCV polyprotein. The protein elicits
XX an immune response in patients infected with HCV and are produced during
XX HCV infection. The present sequence is used in vaccine compositions for
XX preventing HCV infection. It is also used for diagnosing HCV infection
XX and as a target for anti-HCV therapy

XX SQ Sequence 198 AA;
Query Match 94.9%; Score 813.5; DB 3; Length 198;
Best Local Similarity 95.8%; Pred. No. 1.7e-74;
Matches 158; Conservative 1; Mismatches 5; Indels 1; Gaps 1;

QY 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRERLPSGRNLEVD 60
DB 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRERLPSGRNLEVD 60
QY 61 VLSLPRVLGPRAGPGLSPGTLGSPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWR 120
DB 61 VLSLPRVLGPRAGPGLSPGTLGSPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPLRAASPTSGTYRSSAPLLEALPGPWRMASGFWKTA-TMQQ 164
DB 121 SSIPLRAASPTSGTYRSSAPLLEALPGPWRMASGFWKTA-TMQQ 165

RESULT 3
ABU61864
ID ABU61864 standard; protein; 161 AA.
XX AC ABU61864;
XX DT 14-AUG-2003 (first entry)
XX DE HCV core protein frameshift protein p17 #1.
XX KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;
XX KW hepatotropic; overlapping open reading frame; p21c; vaccine.
XX OS Hepatitis C virus type 1.
XX PN US2002076415-A1.
XX 20-JUN-2002.
XX 14-DEC-2000; 2000US-00736959.
XX 14-DEC-1999; 99US-0170835P.
XX (OUJJ/) OU J.
XX (XUZZ/) XU Z.
XX Ou J, Xu Z;
XX WPI; 2003-479366/45.
XX Isolated hepatitis C virus (HCV) proteins formed by expression of
XX overlapping open reading frames in the core protein gene sequence through
XX a frame shifting mechanism, useful for vaccinating against, and detecting
XX HCV infections.
XX Claim 4; Fig 7; 37pp; English.
XX The invention relates to an isolated and purified protein of the
XX hepatitis C virus (HCV) that is formed by expression of an overlapping
XX open reading frame in the core protein gene sequence through an RNA frame
XX shifting mechanism. The protein is termed p17 (the full length, unshifted
XX protein being p21c). Also included are a vaccine (including a DNA
XX vaccine) for immunising a mammal against hepatitis C (producing a
XX protective antibody) comprising at least 1 protein of p17 (or a nucleic
XX acid encoding p17), an anti-viral composition (used to treat hepatitis C)
XX comprising a compound that binds to p17, antibodies directed against an
XX HCV core protein which are elicited by immunising an animal using the
XX partially purified protein p17, a method for analysing an HCV antigen in
XX a sample using the anti-p17 antibodies and detection of anti-HCV
XX antibodies in a sample using the p17 proteins. The HCV p17 and the DNA
XX sequences that encode it may be used as vaccines for immunising patients
XX against HCV infection. The antibodies and the antiviral compound may also
XX be used for treating HCV infections. HCV p17 and the antibodies may also

CC be used in immunoassays for detecting HCV antigens and/or antibodies in
 CC samples for the diagnosis of HCV infections. The present sequence
 CC represents a p17 protein from an HCV strain
 XX
 SQ Sequence 161 AA;
 Query Match 94.0%; Score 806; DB 6; Length 161;
 Best Local Similarity 99.4%; Pred. No. 7.5e-74;
 Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 7 KKKNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRRLPSGRNLEVDVLSLSPR 66
 Db 8 QKNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRRLPSGRNLEVDVLSLSPR 67
 QY 67 LVGPAGPGLSPGTGLGSPSMAMRAAGRDGSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 126
 Db 68 LVGPAGPGLSPGTGLGSPSMAMRAAGRDGSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 127
 QY 127 AASPTSWGTYRSSAPLLEALPGPWRMASGFWKTA 160
 Db 128 AASPTSWGTYRSSAPLLEALPGPWRMASGFWKTA 161
 RESULT 4
 ABU61866
 ID ABU61866 standard; protein; 161 AA.
 XX
 AC ABU61866;
 XX
 DT 23-OCT-2003 (revised)
 DT 14-AUG-2003 (first entry)
 XX
 DE HCV core protein frameshift protein p17 #3.
 XX
 KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;
 KW hepatotropic; overlapping open reading frame; p21c; vaccine.
 XX
 OS Hepatitis C virus; genotype 1a.
 XX
 PN US2002076415-A1.
 XX
 PD 20-JUN-2002.
 XX
 PF 14-DEC-2000; 2000US-00736959.
 XX
 PR 14-DEC-1999; 99US-0170835P.
 XX
 PA (OUJJ/) OU J.
 PA (XUZZ/) XU Z.
 XX
 PI Ou J, Xu Z;
 XX
 DR WPI; 2003-479366/45.
 XX
 PS Isolated hepatitis C virus (HCV) proteins formed by expression of
 PT overlapping open reading frames in the core protein gene sequence through
 PT a frame shifting mechanism, useful for vaccinating against, and detecting
 PT HCV infections.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an isolated and purified protein of the
 CC hepatitis C virus (HCV) that is formed by expression of an overlapping
 CC open reading frame in the core protein gene sequence through an RNA frame
 CC shifting mechanism. The protein is termed p17 (the full length, unshifted
 CC protein being p21c). Also included are a vaccine (including a DNA
 CC vaccine) for immunising a mammal against hepatitis C (producing a
 CC protective antibody) comprising at least 1 protein of p17 (or a nucleic
 CC acid encoding p17), an anti-viral composition (used to treat hepatitis C)
 CC comprising a compound that binds to p17, antibodies directed against an
 CC HCV core protein which are elicited by immunising an animal using the
 CC partially purified protein p17, a method for analysing an HCV antigen in
 CC a sample using the anti-p17 antibodies and detection of anti-HCV

CC antibodies in a sample using the p17 proteins. The HCV p17 and the DNA
 CC sequences that encode it may be used as vaccines for immunising patients
 CC against HCV infection. The antibodies and the antiviral compound may also
 CC be used for treating HCV infections. HCV p17 and the antibodies may also
 CC be used in immunoassays for detecting HCV antigens and/or antibodies in
 CC samples for the diagnosis of HCV infections. The present sequence
 CC represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to
 CC standardise OS field)
 XX
 SQ Sequence 161 AA;
 Query Match 89.6%; Score 768; DB 6; Length 161;
 Best Local Similarity 95.5%; Pred. No. 5.6e-70;
 Matches 147; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 QY 7 KKKNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRRLPSGRNLEVDVLSLSPR 66
 Db 8 QKNTVPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWVCARRRLPSGRNLEVDVLSLSPR 67
 QY 67 LVGPAGPGLSPGTGLGSPSMAMRAAGRDGSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 126
 Db 68 LVGPAGPGLSPGTGLGSPSMAMRAAGRDGSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 127
 QY 127 AASPTSWGTYRSSAPLLEALPGPWRMASGFWKTA 160
 Db 128 AASPTSWGTYRSSAPLLEALPGPWRMASGFWKTA 161
 RESULT 5
 ABU61865
 ID ABU61865 standard; protein; 161 AA.
 XX
 AC ABU61865;
 XX
 DT 23-OCT-2003 (revised)
 DT 14-AUG-2003 (first entry)
 XX
 DE HCV core protein frameshift protein p17 #2.
 XX
 KW HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;
 KW hepatotropic; overlapping open reading frame; p21c; vaccine.
 XX
 OS Hepatitis C virus; genotype 1a.
 XX
 PN US2002076415-A1.
 XX
 PD 20-JUN-2002.
 XX
 PF 14-DEC-2000; 2000US-00736959.
 XX
 PR 14-DEC-1999; 99US-0170835P.
 XX
 PA (OUJJ/) OU J.
 PA (XUZZ/) XU Z.
 XX
 PI Ou J, Xu Z;
 XX
 DR WPI; 2003-479366/45.
 XX
 PS Isolated hepatitis C virus (HCV) proteins formed by expression of
 PT overlapping open reading frames in the core protein gene sequence through
 PT a frame shifting mechanism, useful for vaccinating against, and detecting
 PT HCV infections.
 XX
 PS Claim 9; Page 14; 37pp; English.
 XX
 CC The invention relates to an isolated and purified protein of the
 CC hepatitis C virus (HCV) that is formed by expression of an overlapping
 CC open reading frame in the core protein gene sequence through an RNA frame
 CC shifting mechanism. The protein is termed p17 (the full length, unshifted
 CC protein being p21c). Also included are a vaccine (including a DNA
 CC vaccine) for immunising a mammal against hepatitis C (producing a
 CC protective antibody) comprising at least 1 protein of p17 (or a nucleic
 CC acid encoding p17), an anti-viral composition (used to treat hepatitis C)
 CC comprising a compound that binds to p17, antibodies directed against an
 CC HCV core protein which are elicited by immunising an animal using the
 CC partially purified protein p17, a method for analysing an HCV antigen in
 CC a sample using the anti-p17 antibodies and detection of anti-HCV

QY	127	AASTPSWGTYSRSP	141	Db	8	QRKPNVTPTAAHRTSSRAVRSLSVEFTCCRAGAPGWVCARPCRLQSGRLNVEGDNLSPR	67
Db	128	AASTPSWGTYSRSP	142	QY	67	LVGPRAGPCLSPGTLGSPSMARAAAGRGDSCFLVALGLAGAPQTPGVGRAIWRSSIPLR	126
				Db	68	LANPRAGPCLSPGTLGSPSMATRAWGQDSCHPAALGLIGAPRTPTGVGRAIWRSSIPSR	127
RESULT 13				QY	127	AASTPSWGTYSRSP	141
ABU61873				Db	128	AASTPSWGTYSRSP	142
ID	ABU61873	standard; protein; 143 AA.					
XX	AC	ABU61873;					
XX	DT	23-OCT-2003 (revised)					
XX	DT	14-AUG-2003 (first entry)					
XX	DE	HCV core protein frameshift protein p17 #10.					
XX	KW	HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;					
XX	KW	hepatotropic; overlapping open reading frame; p21c; vaccine.					
XX	OS	Hepatitis C virus; genotype 1b.					
XX	PN	US2002076415-A1.					
XX	PD	20-JUN-2002.					
XX	PF	14-DEC-2000; 2000US-00736959.					
XX	PR	14-DEC-1999; 99US-0170835P.					
XX	PA	(OUJ/) OU J.					
XX	PA	(XUZZ/) XU Z.					
XX	PI	Ou J, Xu Z;					
XX	XX	WPI; 2003-479366/45.					
XX	DR	Isolated hepatitis C virus (HCV) proteins formed by expression of					
XX	PT	overlapping open reading frames in the core protein gene sequence through					
XX	PT	a frame shifting mechanism, useful for vaccinating against, and detecting					
XX	PT	HCV infections.					
XX	XX	Claim 9; Page 15; 37pp; English.					
XX	PS	The invention relates to an isolated and purified protein of the					
XX	CC	hepatitis C virus (HCV) that is formed by expression of an overlapping					
XX	CC	open reading frame in the core protein gene sequence through an RNA frame					
XX	CC	shifting mechanism. The protein is termed p17 (the full length, unshifted					
XX	CC	protein being p21c). Also included are a vaccine (including a DNA					
XX	CC	vaccine) for immunising a mammal against hepatitis C (producing a					
XX	CC	acid encoding p17), an anti-viral composition (used to treat hepatitis C)					
XX	CC	comprising a compound that binds to p17, antibodies directed against an					
XX	CC	HCV core protein which are elicited by immunising an animal using the					
XX	CC	partially purified protein p17, a method for analysing an HCV antigen in					
XX	CC	antibodies in a sample using the p17 proteins. The HCV p17 and the DNA					
XX	CC	sequences that encode it may be used as vaccines for immunising patients					
XX	CC	be used for treating HCV infections. HCV p17 and the antibodies may also					
XX	CC	be used in immunoassays for detecting HCV antigens and/or antibodies in					
XX	CC	samples for the diagnosis of HCV infections. The present sequence					
XX	CC	represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to					
XX	CC	standardise OS field)					
XX	XX	Sequence 143 AA;					
XX	XX	Query Match	63.0%;	Score 540;	DB 6;	Length 143;	
XX	XX	Best Local Similarity	78.5%;	Pred. No. 8.2e-47;			
XX	XX	Matches 106;	Conservative	5;	Mismatches 24;	Indels 0;	Gaps 0;
QY	7	KKKNVTPTAAHRTSSRAVRSLSVEFTCCRAGALDWVCARRRLPGRNLEVDVLSIPR	66				
		:					

Best Local Similarity 77.8%; Pred. No. 8.2e-47;
Matches 105; Conservative 5; Mismatches 25; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSGRNLEVDVSLSPR 66
Db 8 QRPNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSGRNLEVDVSLSPR 67
QY 67 LVGPRAAGPGLSPGTLGSPMAMRAAGGRDSCLPVALGLAGAPQTPGVGRAIWRSSIPLR 126
Db 68 LAGPRAGPGLSPGTLGSPMAMRAAGGRDSCLPVALGLAGAPQTPGVGRAIWRSSIPSH 127
QY 127 AASPTSMGTYSRNSAP 141
Db 128 AASPTSMGTYSRNSAP 142

RESULT 15
ABU61875
ID ABU61875 standard; protein; 143 AA.

XX AC ABU61875;
XX AC
XX DT 23-OCT-2003 (revised)
XX DT 14-AUG-2003 (first entry)
XX HCV core protein frameshift protein p17 #12.
XX DE
XX DE HCV; hepatitis C infection; RNA frameshift; core protein; p17; virucide;
XX KW hepatotropic; overlapping open reading frame; p21c; vaccine.
XX OS
XX OS Hepatitis C virus; genotype 1b.

XX US2002076415-A1.
XX PD 20-JUN-2002.
XX PF 14-DEC-2000; 2000US-00736959.
XX PR 14-DEC-1999; 99US-0170835P.
XX PA (OUJ/) OU J.
XX PA (XUZZ/) XU Z.
XX PI Ou J, Xu Z;
XX WPI; 2003-479366/45.
XX Isolated hepatitis C virus (HCV) proteins formed by expression of
PT overlapping open reading frames in the core protein gene sequence through
PT a frame shifting mechanism, useful for vaccinating against, and detecting
PT HCV infections.

XX Claim 9; Page 15; 37pp; English.
XX The invention relates to an isolated and purified protein of the
CC hepatitis C virus (HCV) that is formed by expression of an overlapping
CC open reading frame in the core protein gene sequence through an RNA frame
CC shifting mechanism. The protein is termed p17 (the full length, unshifted
CC protein being p21c). Also included are a vaccine (including a DNA
CC vaccine) for immunising a mammal against hepatitis C (producing a
CC protective antibody) comprising at least 1 protein of p17 (or a nucleic
CC acid encoding p17), an anti-viral composition (used to treat hepatitis C)
CC comprising a compound that binds to p17, antibodies directed against an
CC HCV core protein which are elicited by immunising an animal using the
CC partially purified protein p17, a method for analysing an HCV antigen in
CC a sample using the anti-p17 antibodies and detection of anti-HCV
CC antibodies in a sample using the p17 proteins. The HCV p17 and the DNA
CC sequences that encode it may be used as vaccines for immunising patients
CC against HCV infection. The antibodies and the antiviral compound may also
CC be used for treating HCV infections. HCV p17 and the antibodies may also
CC be used in immunoassays for detecting HCV antigens and/or antibodies in
CC samples for the diagnosis of HCV infections. The present sequence
CC represents a p17 protein from an HCV strain. (Updated on 23-OCT-2003 to

CC standardise OS field)
XX SQ Sequence 143 AA;
Query Match 62.3%; Score 534; DB 6; Length 143;
Best Local Similarity 77.0%; Pred. No. 3.4e-46;
Matches 104; Conservative 6; Mismatches 25; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSGRNLEVDVSLSPR 66
Db 8 QRPNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSGRNLEVDVSLSPR 67
QY 67 LVGPRAAGPGLSPGTLGSPMAMRAAGGRDSCLPVALGLAGAPQTPGVGRAIWRSSIPLR 126
Db 68 LAGPRAGPGLSPGTLGSPMAMRAAGGRDSCLPVALGLAGAPQTPGVGRAIWRSSIPSH 127
QY 127 AASPTSMGTYSRNSAP 141
Db 128 AASPTSMGTYSRNSAP 142

Search completed: August 25, 2005, 03:33:23
Job time : 167 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:26:38 ; Search time 40 Seconds
(without alignments)
394.489 Million cell updates/sec

Title: US-10-664-038-1
Perfect score: 857
Sequence: 1 ARILNLKKTNTVPTVAHRT.....ALPGPWRMASGFWKATMQQ 164

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	374	43.6	95	2 S44213	core protein - hep
2	97	11.3	363	2 T34931	hypothetical prote
3	96	11.2	479	1 S22542	transcription fact
4	96	11.2	485	1 S22543	transcription fact
5	93.5	10.9	105	2 H72733	hypothetical prote
6	90	10.5	1016	2 T19006	ankyrin related pr
7	89	10.4	479	1 A31753	transcription fact
8	89	10.4	552	2 F75311	ABC transporter, A
9	88.5	10.3	583	1 S22544	transcription fact
10	87.5	10.2	478	1 I47154	transcription fact
11	87	10.2	756	2 D75622	hypothetical prote
12	86	10.0	1617	2 B86483	protein F5U5.15 [1
13	85	9.9	360	2 T37285	collagen dpy-2 - C
14	85	9.9	365	2 A39481	serum response fac
15	85	9.9	825	1 EDBEXD	immediate-early pr
16	85	9.9	1042	1 GCCHIS	collagen alpha 1(I
17	84	9.8	143	2 E72699	hypothetical prote
18	84	9.8	152	2 T34649	hypothetical prote
19	84	9.8	1113	2 T14360	period protein Per
20	84	9.8	1115	2 T13955	period protein Per
21	83.5	9.7	387	1 DYHUD4	dopamine receptor
22	83	9.7	229	2 C70753	hypothetical prote
23	83	9.7	784	2 A26601	elastin precursor
24	82.5	9.6	419	2 G70602	hypothetical prote
25	82.5	9.6	538	2 T27156	hypothetical prote
26	82.5	9.6	1573	2 S01845	DNA (cytosine-5-) -
27	81.5	9.5	395	1 KISNG	galactokinase (EC
28	81.5	9.5	671	1 CGRTIS	collagen alpha 1(I
29	81.5	9.5	1442	2 T42607	transcription acti

30 81 9.5 290 2 T24586 hypothetical prote
31 81 9.5 439 1 A48099 transcription fact
32 81 9.5 441 2 A83370 probable MFS trans
33 81 9.5 673 1 CGB06C collagen alpha 1(I
34 81 9.5 1691 1 CGHU6B collagen alpha 6(I
35 81 9.5 1732 2 T43026 probable DNA-direc
36 80.5 9.4 494 2 S39607 transcription fact
37 80.5 9.4 542 2 S39608 transcription fact
38 80.5 9.4 631 1 A36749 transcription fact
39 80.5 9.4 756 2 T00367 hypothetical prote
40 80.5 9.4 1453 2 S21626 collagen alpha 1(I
41 80 9.3 538 2 S57459 hook-containing pr
42 80 9.3 897 1 A39255 cytokine receptor
43 80 9.3 900 2 B70694 probable INF - My
44 79.5 9.3 409 2 A70932 probable PPE prote
45 79.5 9.3 1102 2 T02782 probable relaxase

ALIGNMENTS

RESULT 1

S44213
core protein - hepatitis C virus
C/Species: hepatitis C virus
C/Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C/Accession: S44213
R/Feucht, H.H.
submitted to the EMBL Data Library, April 1994
A/Reference number: S44213
A/Accession: S44213
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-95 <FEU>
A/Cross-references: UNIPROT:Q68874; EMBL:X78950; NID:g475172; PIDN:CAA55547.1; PID:g860

Query Match 43.6%; Score 374; DB 2; Length 95;
Best Local Similarity 78.7%; Pred. No. 1.6e-25;
Matches 74; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 11 NVTPTVAHRTSSSRVAVRSLVFTCCRAGALDWVCARRRLPSGRNLVDVSLPRLVGP 70
Db 2 NVTPTVAHRTSSSRVAVRSLVFTCCRAGAPGWVCARPGRLPSGRNLVEGDNLSPRLASP 61

Qy 71 RAGPGLSPGTLGFSMAMRAAGRGDSCLPVALGL 104
Db 62 RAGPGLSPGTLGFSMAMRVWGGQDGSCHTAPGL 95

RESULT 2

T34931
hypothetical protein SC3F9.09 SC3F9.09 - Streptomyces coelicolor
C/Species: Streptomyces coelicolor
C/Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 09-Jul-2004
C/Accession: T34931
R/Seeger, K.J.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, June 1998
A/Reference number: Z21562
A/Accession: T34931
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-363 <SEE>
A/Cross-references: UNIPROT:O69949; EMBL:AL023862; PIDN:CAA19632.1; GSPDB:GN00070; SCOE1
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB:SC3F9.09

Query Match 11.3%; Score 97; DB 2; Length 363;
Best Local Similarity 30.6%; Pred. No. 0.47;
Matches 44; Conservative 14; Mismatches 40; Indels 46; Gaps 8;

Qy 39 GALDWVCARRRRLPSGRNLVDVSLPRLVGPACGLSPGTLGFSMAMRAAGRGDSC 97
Db 39 GALTVAHRTSSSRVAVRSLVFTCCRAGALDWVCARRRLPSGRNLVEGDNLSPRLASP 61

Db 244 GALTGLVARRRRLPAGRR-----AAAAGLGGSTPAPAVT-----DA 281

Qy 98 LPVALGLAGA-----PQTGCVGRAI--WVRSSIPLRAASPTSGTYRSGAPL 142

Db 282 LAALGAAYATAARPYFHAALNSPPLTRAVGGIRATIPLOAALAARSGA--SATSL 339

Qy 143 LEALPGWRMASGFW--KTATMQQ 164

Db 340 LVAALAP-----AGRWFAKRSAMRK 359

RESULT 3

S22542

transcription factor Oct-2, splice form Oct-2.2 - mouse

N;Alternate names: NF-A2; OTF-2

C;Species: Mus musculus (house mouse)

C;Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S22542; S50017; S47215

R;Wirth, T.; Priess, A.; Annweiler, A.; Zwillling, S.; Oeler, B.

Nucleic Acids Res. 19, 43-51, 1991

A;Title: Multiple Oct2 isoforms are generated by alternative splicing.

A;Reference number: S22539; MUID:91187647; PMID:2011512

A;Accession: S22542

A;Status: preliminary

A;Molecule type: mRNA

A;Cross-references: UNIPROT:Q00196; EMBL:X57937; NID:G53483; PIDN:CAA41005.1; PID:G53484

R;Matsumoto, K.; Clay, O.; Kuenzler, P.; Georgiev, O.; Urbanek, P.; Schaffner, W.

A;Title: Short introns interrupting the Oct-2 POU domain may prevent recombination between

A;Reference number: S50016; MUID:95194574; PMID:7888080

A;Accession: S50017

A;Status: translation not shown

A;Molecule type: DNA

A;Residues: 168-377 <MAT>

A;Cross-references: EMBL:X81031; NID:G534039; PIDN:CAA56934.1; PID:G536781

C;Comment: This protein is a tissue-specific transcriptional trans-activator.

C;Genetics:

A;Gene: Oct-2

A;Map position: 7

A;Introns: 221/3; 269/1; 318/3

C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

F;202-269/Domain: POU domain homology <POU>

F;298-354/Domain: homeobox homology <HOX>

Query Match 11.2%; Score 96; DB 1; Length 479;

Best Local Similarity 25.4%; Pred. No. 0.76; Mismatches 28; Indels 64; Gaps 6;

Matches 34; Conservative 8

Qy 43 WVCARRR-----LPS-GRNLEVDVLSPLVGPVPRAGPGLSP-----78

Db 344 WFCNRQKEKRINPCSAAPMLPSPGK---PTSYPHLVTPQGGAGTLPQSSSLSTT 399

Qy 79 -----GTGFSNMRAGRGDSCLPV-----ALGL 104

Db 400 VTTLSAVGTLPSPRTAGGGGGGGRALPLNSIPVTPPPATNTNTPSPQSGHSAIGL 459

Qy 105 AGAPQTPGVGRAIW 118

Db 460 SGL--NPSAGPGLW 471

RESULT 4

S22543

transcription factor Oct-2 splice form Oct-2.3 - mouse

N;Alternate names: NF-A2 protein; OTF-2 protein

N;Contains: transcription factor Oct-2 splice form Oct-2.1; transcription factor Oct-2

C;Species: Mus musculus (house mouse)

C;Date: 22-Nov-1993 #sequence revision 10-Nov-1995 #text_change 09-Jul-2004

C;Accession: S22543; S22541; S22540; JH0596; I48698; S47215

R;Wirth, T.; Priess, A.; Annweiler, A.; Zwillling, S.; Oeler, B.

Nucleic Acids Res. 19, 43-51, 1991

A;Title: Multiple Oct2 isoforms are generated by alternative splicing.

A;Reference number: S22539; MUID:91187647; PMID:2011512

A;Accession: S22543

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-485 <WIR>

A;Cross-references: UNIPROT:Q00196; EMBL:X57938; NID:G53485; PIDN:CAA41006.1; PID:G53481

A;Note: splice form Oct-2.3

A;Accession: S22541

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-62,85-485 <WI3>

A;Cross-references: EMBL:X57936; NID:G53481; PIDN:CAA41004.1; PID:G53482

A;Note: splice form Oct-2.1

A;Accession: S22540

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-62,124-485 <WI4>

A;Cross-references: EMBL:X57941; NID:G53491; PIDN:CAA41009.1; PID:G53492

R;Stoykova, A.S.; Sterrer, S.; Erselius, J.R.; Hatzopoulos, A.K.; Gruss, P.

Neuron 8, 541-558, 1992

A;Title: Mini-Oct and Oct-2c: two novel, functionally diverse murine Oct-2 gene products

A;Reference number: JH0596; MUID:92198662; PMID:1550877

A;Accession: JH0596

A;Molecule type: mRNA

A;Residues: 1-62,85-430,'A',432,'P',434-473 <STO>

A;Experimental source: brain

A;Note: splice form Oct-2c

R;Stepchenko, A.G.

Dokl. Akad. Nauk SSSR 325, 175-178, 1992

A;Title: Interaction of Oct-binding transcription factors with a large series of 'nonc

A;Reference number: I48698

A;Accession: I48698

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 'PHP',189-430,'A',432,'P',434-464,'W',466-485 <RES>

A;Cross-references: EMBL:X57089; NID:G53493; PIDN:CAA40369.1; PID:G53494

C;Comment: This protein is a tissue-specific transcriptional trans-activator.

C;Genetics:

A;Gene: Oct-2

A;Map position: 7

C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation

F;208-275/Domain: POU domain homology <POU>

F;304-360/Domain: homeobox homology <HOX>

Query Match 11.2%; Score 96; DB 1; Length 485;

Best Local Similarity 25.4%; Pred. No. 0.77; Mismatches 28; Indels 64; Gaps 6;

Matches 34; Conservative 8

Qy 43 WVCARRR-----LPS-GRNLEVDVLSPLVGPVPRAGPGLSP-----78

Db 350 WFCNRQKEKRINPCSAAPMLPSPGK---PTSYPHLVTPQGGAGTLPQSSSLSTT 405

Qy 79 -----GTGFSNMRAGRGDSCLPV-----ALGL 104

Db 406 VTTLSAVGTLPSPRTAGGGGGGGRALPLNSIPVTPPPATNTNTPSPQSGHSAIGL 465

Qy 105 AGAPQTPGVGRAIW 118

Db 466 SGL--NPSAGPGLW 477

RESULT 5

H72733

hypothetical protein APE0408 - Aeropyrum pernix (strain K1)

C;Species: Aeropyrum pernix

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: H72733

R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain 122
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-105 <KAW>
A;Cross-references: UNIPROT:Q9YF32; DDBJ:AP000059; NID:G51033911; PIDN:BAA79364.1; PID:G51033911
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0408

Query Match 10.9%; Score 93.5; DB 2; Length 105;
Best Local Similarity 33.7%; Pred. No. 0.27;
Matches 34; Conservative 11; Mismatches 33; Indels 23; Gaps 6;

QY 64 SPRLVG---PRAGPLSPGLTSPGMANRAAG-----GRDQSCPLVA-----LGLAGA 107
Db 5 SIRLVSTLPRASP---PSTLSPEFS--GSLKWNFLSLGRGACIPLSTNILLGSTSA 59
QY 108 PQTGPGVGRAIWVRSSIPLRASPTSMCTYRSSAPLLEALPG 148
Db 60 TFSNMRGLWTSYIVPRRPGSS--TLRGAPVPTTSPG 98

RESULT 6
T19006
ankyrin related protein C06C3.1 - Caenorhabditis elegans
N;Contains: myosin-light-chain-phosphatase (EC 3.1.3.53)
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19006; T22086
R;Barks, M.
submitted to the EMBL Data Library, August 1994
A;Reference number: Z19058
A;Accession: T19006
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1016 <WIL>
A;Cross-references: UNIPROT:Q17718; EMBL:Z36719; PIDN:CAA85318.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone C06C3
R;Matthews, P.
submitted to the EMBL Data Library, January 1995
A;Reference number: Z19510
A;Accession: T22086
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1016 <W12>
A;Cross-references: EMBL:Z47809; PIDN:CAA87782.1; GSPDB:GN00020; CESP:C06C3.1
A;Experimental source: clone F42A8
C;Genetics:
A;Gene: CESP.C06C3.1
A;Map position: 3
A;Introns: 27/3; 94/3; 279/3; 352/2; 462/2; 523/3; 569/2; 657/2; 718/3; 766/3; 833/3; 893/3
C;Keywords: phosphoric monoester hydrolase

Query Match 10.5%; Score 90; DB 2; Length 1016;
Best Local Similarity 22.3%; Pred. No. 5.3;
Matches 39; Conservative 22; Mismatches 60; Indels 54; Gaps 6;

QY 6 LKKTNTVPTVAHRTSSRVAV-RSLVEFTCCRAGALDWYCARRRLPSCGRNLEVDVSL 64
Db 439 LKNGKEISPLRSETTSSRSISITRSLDGYT-----DRSSSGRETSAEWS-- 482
QY 65 PRUVGPAGPLSPGLTILG-----PSMAMRAAG-----RD 94
Db 483 -----EAASSASTGTSSSRFTSSPTSSQSRGAAGSVHTTPRSMENSSVSGSDQNV 536
QY 95 GSCPLVALGLAGAPQTPGVGRATWVRSSIPLRASPTSMCTYRSSAPLLEALPG 149
Db 537 SATIPI-VPLSAPPKAVHQSPSWINRGVPLSSRSSTSSVTRSSSTPVSSEILSP 590

RESULT 7
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix strain 122
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: H72733
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-105 <KAW>
A;Cross-references: UNIPROT:Q9YF32; DDBJ:AP000059; NID:G51033911; PIDN:BAA79364.1; PID:G51033911
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0408

transcription factor Oct-2 - human
N;Alternate names: NF-A2; OTF-2
C;Species: Homo sapiens (man)
C;Date: 07-Jun-1990 #sequence_revision 26-Apr-1996 #text_change 09-Jul-2004
C;Accession: A31753; C31753; S06452; A31213; S47513; S06453; S50016; S12535
R;Clerc, R.G.; Corcoran, L.M.; LeBowitz, J.H.; Baltimore, D.; Sharp, P.A.
Genes Dev. 2, 1570-1581, 1988
A;Title: The B-cell-specific Oct-2 protein contains POU box- and homeo box-type domains
A;Reference number: A91621; MUID:89107992; PMID:3265124
A;Accession: A31753
A;Molecule type: mRNA
A;Residues: 1-467 <CLE>
A;Cross-references: UNIPROT:P09086; EMBL:X53468; NID:G35128; PIDN:CAA37562.1; PID:G3512
A;Accession: C31753
A;Molecule type: mRNA
A;Residues: 450-479 <CL2>
A;Cross-references: EMBL:X53469; NID:G35131; PIDN:CAA37565.1; PID:G35132; EMBL:Y00227
R;Mueller, M.M.; Ruppert, S.; Schaffner, W.; Matthias, P.
Nature 336, 544-551, 1988
A;Title: A cloned octamer transcription factor stimulates transcription from lymphoid-
A;Reference number: S06452; MUID:89070674; PMID:2904653
A;Accession: S06452
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 2-479 <MUE>
A;Cross-references: EMBL:X13809
A;Note: It is uncertain whether Met-6 or Met-13 is the initiator or whether translation
R;Ko, H.S.; Fast, P.; McBride, W.; Staudt, L.M.
Cell 55, 135-144, 1988
A;Title: A human protein specific for the immunoglobulin octamer DNA motif contains a fi
A;Reference number: A31213; MUID:89030042; PMID:2901913
A;Accession: A31213
A;Molecule type: mRNA
A;Residues: 255-402 <KOH>
A;Cross-references: GB:M22596; NID:G727159; PIDN:AAA64232.1; PID:G727160
R;Natsuo, K.; Clay, O.; Kuenzler, P.; Georgiev, O.; Urbanek, P.; Schaffner, W.
submitted to the EMBL Data Library, August 1994
A;Reference number: S47215
A;Accession: S47513
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 168-377 <MAT>
A;Cross-references: EMBL:X81030
R;Scheiderer, C.; Cromlish, J.A.; Gerster, T.; Kawakami, K.; Balmaceda, C.G.; Currie, I.
Nature 336, 551-557, 1988
A;Title: A human lymphoid-specific transcription factor that activates immunoglobulin g
A;Reference number: S06453; MUID:89070675; PMID:2904654
A;Accession: S06453
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-167,184-479 <SCH>
A;Cross-references: EMBL:X13810; NID:G35166; PIDN:CAA32040.1; PID:G35167
R;Natsuo, K.; Clay, O.; Kuenzler, P.; Georgiev, O.; Urbanek, P.; Schaffner, W.
Biol. Chem. Hoppe-Seyler 375, 675-683, 1994
A;Title: Short introns interrupting the Oct-2 POU domain may prevent recombination betw
A;Reference number: S50016; MUID:95194574; PMID:7888080
A;Accession: S50016
A;Status: preliminary; translation not shown
A;Molecule type: DNA
A;Residues: 168-377 <MA2>
A;Cross-references: EMBL:X81030; NID:G532054; PIDN:CAA56933.1; PID:G4379032
R;Mueller-Immervogel, M.M.; Schaffner, W.; Matthias, P.
EMBO J. 9, 1625-1634, 1990
A;Title: Transcription factor Oct-2A contains functionally redundant activating domains
A;Reference number: S12535; MUID:90228363; PMID:2328728
A;Accession: S12535
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 99-161 <MUT2>
C;Comment: This protein is a tissue-specific transcriptional trans-activator. In humans
C;Genetics:
A;Gene: GDB:POU2F2; OCT2; OTF2

A;Cross-references: GDB:120255; OMIM:164176

A;Map position: 19pter-19qter

A;Introns: 221/3; 269/1; 316/3

C;Superfamily: transcribed factor Oct-2; homeobox homology; POU domain homology

C;Keywords: alternative splicing; B-cell; DNA binding; homeobox; nucleus; transcription

F;202-269/Domain: POU domain homology <POU>

F;298-354/Domain: homeobox homology <HOX>

Query Match 10.4%; Score 89; DB 1; Length 479;

Best Local Similarity 23.9%; Pred. No. 3;

Matches 32; Conservative 10; Mismatches 28; Indels 64; Gaps 6;

Qy 43 WVCARRER-----LPS-GRNLEVDVSLPRLVGPAGPGLSP-----78

Db 344 WFCNRQKEKRINPCSAAPMLPSGK---PASYSPEHVTPOGGAGTLPQSASSLSSTT 399

Qy 79 -----GTLGSPMAMRAAGRGDSCLPV-----ALGL 104

Db 400 VTTLSAVGTLPHTSRTAGGGGGGAAPPLNSIPSVTPPTTNTNPSPOGSHSAIGL 459

Qy 105 AGAPQTPGVGRAIW 118

Db 460 SGL--NPSTGGLW 471

RESULT 8

F75311

ABC transporter, ATP-binding protein - Deinococcus radiodurans (strain R1)

C;Species: Deinococcus radiodurans

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: F75311

R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: F75311

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-552 <WHI>

A;Cross-references: UNIPROT:Q9RS9; GB:AB002048; GB:AB000513; NID:96459929; PIDN:AAF1168

A;Experimental source: strain R1

C;Genetics:

A;Gene: DR2145

A;Map position: 1

Query Match 10.4%; Score 89; DB 2; Length 552;

Best Local Similarity 29.5%; Pred. No. 3.5;

Matches 49; Conservative 10; Mismatches 55; Indels 52; Gaps 9;

Qy 18 HRTSSSRVAVSLVEFTCCAGALDWTVCARRERLPGRNLEVDVSLPRLVGPAGP-- 75

Db 119 HAAANARRA-----SCARERHPD-VCAARR-----TGSAPFRAGAAYPE----RGAGRR 162

Qy 76 --LSPGTLGSPMAMRAAGRGDSCLPVALGAGAPQPGVGRAIW-----VRSSIP 124

Db 163 GVLSSHSLGTDRAARRAAGRGGR--TPAAGV-SDARTTRPVGRTDTRRGALAGARLSHP 220

Qy 125 LRAASPTSWGTYRSSA-----PLLEALPGP 149

Db 221 LAAAPGDGAPRDPARAPAAAPAAARIHRWPERSGPRPRRRAAPGP 266

RESULT 9

S22544

transcription factor Oct-2, splice form Oct-2.5 - mouse

N;Alternate names: NF-A2; OTF-2; transcription factor Oct-2b

C;Species: Mus musculus (house mouse)

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004

C;Accession: S22544; A60085; S09240

R;Wirth, T.; Pries, A.; Annweiler, A.; Zwillig, S.; Oeler, B.

Nucleic Acids Res. 19, 43-51, 1991

Query Match 10.3%; Score 88.5; DB 1; Length 583;

Best Local Similarity 22.9%; Pred. No. 4.1;

Matches 47; Conservative 14; Mismatches 41; Indels 103; Gaps 9;

Qy 43 WVCARRER-----LPS-GRNLEVDVSLPRLVGPAGPGLSP-----78

Db 328 WFCNRQKEKRINPCSAAPMLPSGK---PTSYPHLVTPQGGAGTLPQSASSLSSTT 383

Qy 79 -----GTLGSPMAMRAAGRGDSCLPV-----ALGL 104

Db 384 VTTLSAVGTLPHTSRTAGGGGGGAAPPLNSIPSVTPPTTNTNPSPOGSHSAIGL 443

Qy 105 AGAPQTPG--VG-----RAIWRSSIPLRA-----ASP 130

Db 444 SGLNPSAGSTWGLSGSLSPALMSNNPLATICALASGGTLPILTSLDGSGNLVLAGAAP 503

RESULT 10

I47154

transcription factor Oct-2 - pig

N;Alternate names: NF-A2; OTF-2

C;Species: Sus scrofa domestica (domestic pig)

C;Date: 21-Feb-1997 #sequence_revision 21-Feb-1997 #text_change 09-Jul-2004

C;Accession: I47154

R;Tuggle, C.K.; Helm, J.; Rothschild, M.F.

Anim. Genet. 25, 141-145, 1994

A;Title: Cloning, sequencing and restriction fragment length polymorphism analysis of a

A;Reference number: I47154; MUID:95030552; PMID:7943947

A;Accession: I47154

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-478 <TUG>

A;Cross-references: UNIPROT:Q29013; EMBL:U00794; NID:9451313; PIDN:AAA80148.1; PID:9451;

C;Comment: This protein is a tissue-specific transcriptional trans-activator.

C;Genetics:

A;Gene: OCT2

C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

A;Title: Multiple Oct2 isoforms are generated by alternative splicing.

A;Reference number: S22539; MUID:91187647; PMID:2011512

A;Accession: S22544

A;Molecule type: mRNA

A;Residues: 1-583 <WR>

A;Cross-references: UNIPROT:Q00196; EMBL:X57940; NID:953489; PIDN:CAA1008.1; PID:953491

R;Hatzopoulos, A.K.; Stoykova, A.S.; Erselius, J.R.; Goulding, M.; Neuman, T.; Gruss, P.

Development 109, 349-362, 1990

A;Title: Structure and expression of the mouse Oct2a and Oct2b, two differentially splic

A;Reference number: A60085; MUID:90382251; PMID:1976089

A;Accession: A60085

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-408, 'A', 410, 'P', 412-583 <HAT>

A;Cross-references: EMBL:X53654; NID:9288173; PIDN:CAA37702.1; PID:9288174

R;Goldborough, A.; Ashworth, A.; Willison, K.

Nucleic Acids Res. 18, 1634, 1990

A;Title: Cloning and sequencing of POU-boxes expressed in mouse testis.

A;Reference number: S09237; MUID:90221898; PMID:1970171

A;Accession: S09240

A;Molecule type: DNA

A;Residues: 203-325 <GOL>

A;Cross-references: EMBL:X51961; NID:953495; PIDN:CAA36220.1; PID:9930194

C;Comment: This protein is a tissue-specific transcriptional trans-activator.

C;Genetics:

A;Gene: Oct-2

A;Map position: 7

C;Superfamily: transcription factor Oct-2; homeobox homology; POU domain homology

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulat

F;186-253/Domain: POU domain homology <POU>

F;282-338/Domain: homeobox homology <HOX>

C;Keywords: alternative splicing; DNA binding; homeobox; nucleus; transcription regulation
F;202-269/Domain; POU domain homology <POU>
F;298-354/Domain: homeobox homology <HOX>

Query Match 10.2%; Score 87.5; DB 1; Length 478;
Best Local Similarity 26.7%; Pred. No. 4.1;
Matches 44; Conservative 9; Mismatches 37; Indels 75; Gaps 10;
Qy 43 WVCARRER-----LPS-GRNLEVDVSLPRLVGPAGPGLSP----- 78
Db 344 WFCNRQKEKINPCSAAPMLSPGK-----PASYSPLVTPQGGAGTLPQSAASSLSTT 399
Qy 79 -----GTLGPMAMRAAGRGDSCLPVALGAGAPQTPGVGRAIWRSSIPRLRAASP 130
Db 400 VTTLSAVGTLLHPS-----RTAGG-----GAAGGAAPPL-----NSIP--SVTP 437
Qy 131 TSWGTYRSSAPL-----LEALPGPMRWASGFWKATMQ 163
Db 438 PPAATNTSNPQSGHSAIGLGNLPSTGP-----GLWNPAPYQ 477

RESULT 11
D75622
hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: D75622
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: D75622
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-756 <WHI>
A;Cross-references: UNIPROT:Q9RZSS; GB:AE001826; NID:G6460827; PIDN:AAF12630.1; PID:G646
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRB0039
A;Map position: megaplasmid
A;Genome: plasmid
A;Note: plasmid NP1

Query Match 10.2%; Score 87; DB 2; Length 756;
Best Local Similarity 25.6%; Pred. No. 7.2;
Matches 46; Conservative 14; Mismatches 62; Indels 58; Gaps 9;
Qy 10 TNYTPTVAHRTSSSRVAVRSVLVEFTCCRAGALDWVCARRERLPSSGRNLEVDVSLPRLVG 69
Db 286 SNATADLIRRVETQRPGLRK-----TVSKVGAPDW--APGSVATSGDELYTLSTVTPYQA 339
Qy 70 PRAG-PGLSPGTLGPMAMRAAGRGDSCLPVALGAGAPOT-----PGVGRATWIR 120
Db 340 PLAGVQVLDPLPAGCTEFVSADSG-----ALLEGAASAAQAVANTGLDLPAGARTITLIR 392
Qy 121 -----SSPL-----RAASPTSGTYSAPL-----EALPG 148
Db 393 VRGRDVRDQELNRVFLTSTSELPAPLHSAASAVVWNT-----APLLSKTLDRDRAAPG 448

RESULT 12
B86483
Protein FSJ5.15 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: B86483
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: B86483
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1617 <STO>
A;Cross-references: UNIPROT:Q9SKV5; GB:AE005172; NID:G6598587; PIDN:AAF18642.1; GSPDB:G
C;Genetics:
A;Gene: FSJ5.15
A;Map position: 1

Query Match 10.0%; Score 86; DB 2; Length 1617;
Best Local Similarity 26.8%; Pred. No. 19;
Matches 34; Conservative 16; Mismatches 55; Indels 22; Gaps 4;
Qy 54 GRNLEVDVSLPRL--VGPAGPGL-----SPTGLGPMAMRAAGRGDSCLPVA 101
Db 82 GRSATIEQSVGSQLEFVEPGVGDGLGADIAGATGVGAGGAGSVGVHAGAKDPGLGVA 141
Qy 102 LGLAGAPQTPGVGRAIWR-----SSIPLRASPTSWGTYRSSAPLLEALPGFWRMAS 154
Db 142 ---AGGAQVPEVGLAGLLRQLRLPGVVPVHAPVAPRAEVOQRAAAVVEVPSYLRME 198
Qy 155 GFWKAT 161
Db 199 QLQRIQT 205

RESULT 13
T37285
collagen dpy-2 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C;Accession: T37285
R;Levy, A.D.; Yang, J.; Kramer, J.M.
Mol. Biol. Cell 4, 803-817, 1993
A;Title: Molecular and genetic analyses of the Caenorhabditis elegans dpy-2 and dpy-10
A;Reference number: Z21666; MUID:94060446; PMID:8241567
A;Accession: T37285
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-360 <LEV>
A;Cross-references: EMBL:L12706; NID:G289665; PIDN:AAAL7398.1; PID:G467810
A;Note: a variety of molecular defects in these collagens can result in severe morpholo;
C;Genetics:
A;Gene: dpy-2
A;Introns: 18/3; 68/3; 118/1; 325/3
C;Superfamily: collagen alpha 2(I) chain; fibrillar collagen carboxyl-terminal homology

Query Match 9.9%; Score 85; DB 2; Length 360;
Best Local Similarity 24.4%; Pred. No. 5;
Matches 47; Conservative 10; Mismatches 56; Indels 80; Gaps 9;
Qy 32 EFTCCRAGALDW--VCARRERLPSSGRNL-----EVDVSLSP 65
Db 63 DFECQASANDLETSMVSVEGLLRGNVTKRAAGYGHYNPMLAADSPOFQECFASCIP 122
Qy 66 RLVGPRAGPGLS--PGTLGPMAMRAAGRGDS-----CLPVALGLAGA 107
Db 123 GERGPSGDSGLPALPGAPGD---GAPGRPGTTNASCIPERVFEPFPCPLPCPGPAGV 178
Qy 108 PQTPGV-----GRAIWRSSIPLRASPTSWGTY---RSSAPLLE----- 144
Db 179 PGHPGFPDGEYIGIGRP--GSDGMPGKPGDGLAGIPGPGSGSIGDKRTPEARVI 236
Qy 145 -----ALPGPW 150
Db 237 PGPPGSGGLPGPW 249

RESULT 14

A39481
serum response factor-related protein 2 - human
N/Alternate names: myocyte-specific enhancer factor xmf2; RSRF2
C/Species: Homo sapiens (man)
C/Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 09-Jul-2004
C/Accession: A39481; S25832; S24468
R/Pollack, R.; Treisman, R.
Genes Dev. 5, 2327-2341, 1991
A/Title: Human SRF-related proteins: DNA-binding properties and potential regulatory tar
A/Reference number: A39481; MUID:92084105; PMID:1748287
A/Accession: A39481
A/Molecule type: mRNA
A/Residues: 1-365 <POL>
A/Cross-references: UNIPROT:Q02080; EMBL:X63380; NID:G36168; PIDN:CAA44978.1; PID:G36168
R/Yu, Y.T.; Breitbart, R.E.; Smoot, L.B.; Lee, Y.; Mahdavi, V.; Nadal-Ginard, B.
Genes Dev. 6, 1783-1798, 1992
A/Title: Human myocyte-specific enhancer factor 2 comprises a group of tissue-restricted
A/Reference number: S25830; MUID:92387551; PMID:1516833
A/Accession: S25832
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-365 <YU>
A/Cross-references: EMBL:X68502; NID:G37991; PIDN:CAA48515.1; PID:G37992
C/Keywords: DNA binding; homodimer; transcription factor
F/2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 9.9%; Score 85; DB 2; Length 365;
Best Local Similarity 25.2%; Pred. No. 5.1;
Matches 36; Conservative 10; Mismatches 47; Indels 50; Gaps 7;
QY 47 RRRLPSGRNLEVDVLSPLRVGPRGGLSPGTLGFSMAARAAGRGDGLPVALG--L 104
DB 207 RRSLLPGG-----LAGPRG-----LNTSRSLYS--GLQNFCSATFGPPL 245
QY 105 AGAPQTG-----VGRAIWVR-----SSIPLRASPTSGWTVRSAPLLEA 145
DB 246 GSFPFLGPPVGAEAWARVPQAPAPRRPPQSSLSASLRPPGAPATFLRSPSPCS 305
QY 146 LQGPWRNMGCF-----WKTA 160
DB 306 SPGPWQSLGLGPPCAGCPMPTA 328

RESULT 15

EDBEXD
Immediate-early protein RL2 - human herpesvirus 2 (strain HG52)
N/Alternate names: RL2 protein
C/Species: human herpesvirus 2
A/Note: host Homo sapiens (man)
C/Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C/Accession: JQ1501
R/McGeoch, D.J.; Cunningham, C.; McIntyre, G.; Dolan, A.
J. Gen. Virol. 72, 3057-3075, 1991
A/Title: Comparative sequence analysis of the long repeat regions and adjoining parts of
A/Reference number: JQ1494; MUID:92113549; PMID:1662697
A/Accession: JQ1501
A/Molecule type: DNA
A/Residues: 1-825 <MCG>
A/Cross-references: UNIPROT:P28284; GB:D10471; DDBJ:D01128; NID:G221784; PIDN:BAA23427.1
C/Genetics:
A/Gene: RL2
A/Introns: 25/3; 252/1
C/Superfamily: herpesvirus immediate-early protein IE110; RING finger homology
C/Keywords: DNA binding; immediate-early protein; tandem repeat; transcription regulatio
F/122-172/Domain: RING finger homology <RNG>
F/126-166/Region: zinc finger C3HC4 motif
F/589-623/Region: 5-residue repeats (A-S-S-S)

Query Match 9.9%; Score 85; DB 1; Length 825;
Best Local Similarity 27.0%; Pred. No. 12;
Matches 40; Conservative 11; Mismatches 63; Indels 34; Gaps 5;

QY 14 PTVAHRTSSRV--AVRSLVEFTCCAGALDWYCARRERLPGRNLEVDVLSPLRVGPR 71
DB 446 PVDAAHRAPRSRMTQAQTDTQAQSLGRAGATD---ARGSGGP-----GAE 486
QY 72 AGPGLSPGTLGFSMAARA-----GGRDGSLPVALGLAGAPQTGPGVGRAIWVR 120
DB 487 GSGFVPRGTTNTGCAAPHAAGAARPRKRGSDSGPAASSSSAAPRSPPLAPQGVGAK 546
QY 121 SSIPLRASPTSGWTVRSAPLLEALPG 148
DB 547 RAAPRRA--PDSDSGRDGRHGGLAPASAG 572

Search completed: August 25, 2005, 03:37:08
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:18:18 ; Search time 171 Seconds
(without alignments)

491.117 Million cell updates/sec

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARILNLKKKTNTVPTVAHRT.....ALFPGWRMASGFWKTATMQQ 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03.:

1: uniprot_sprot.:

2: uniprot_trembl.:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	577	67.3	125	2	Q66VR7 hepatitis c
2	557	65.0	125	2	Q66VM9 hepatitis c
3	556	64.9	125	2	Q66VK1 hepatitis c
4	432.5	50.5	126	2	Q66VN1 hepatitis c
5	403	47.0	122	2	Q00687 hepatitis c
6	374	43.6	95	2	Q68874 hepatitis c
7	350.5	40.9	134	2	Q706D2 hepatitis c
8	346	40.4	102	2	Q68361 hepatitis c
9	336	39.2	102	2	Q68358 hepatitis c
10	331.5	38.7	102	2	Q68356 hepatitis c
11	319	37.2	102	2	Q68359 hepatitis c
12	305	35.6	100	2	Q68360 hepatitis c
13	301	35.1	71	2	Q91AW2 hepatitis c
14	296	34.5	71	2	Q91AW1 hepatitis c
15	287	33.5	102	2	Q68365 hepatitis c
16	285	33.3	102	2	Q68366 hepatitis c
17	262	30.6	102	2	Q68363 hepatitis c
18	262	30.6	102	2	Q68368 hepatitis c
19	254	29.6	102	2	Q68367 hepatitis c
20	225	25.3	80	2	Q81289 hepatitis c
21	220.5	25.7	119	2	Q66686 hepatitis c
22	148	17.3	53	2	Q91KM9 hepatitis c
23	142.5	16.6	125	2	Q66VT0 hepatitis c
24	142.5	16.6	125	2	Q66VV1 hepatitis c
25	142.5	16.6	125	2	Q66VV3 hepatitis c
26	137.5	16.0	125	2	Q66VP7 hepatitis c
27	137.5	16.0	125	2	Q66VQ5 hepatitis c
28	123.5	14.4	210	2	Q6NVL1 homo sapien
29	109.5	12.8	108	2	F87760 hepatitis c
30	109.5	12.8	108	2	F87761 hepatitis c
31	102.5	12.0	514	1	LWA_ANTEL anthopleura

RESULT 1

Q66VR7 ID Q66VR7 PRELIMINARY; PRT; 125 AA.
AC Q66VR7;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15302945; DOI=10.1099/vir.0.80071-0;
RA Deforges S., Eylaheav A., Perret M., Sodoyer M., Pouzol S.,
RA Scoazec J.Y., Bonnaud B., Diaz O., Paranhos-Baccala G., Lotteau V.,
RA Andre P.;
RT "Expression of hepatitis C virus proteins in epithelial intestinal
RT cells in vivo."
RL J. Gen. Virol. 85:2515-2523 (2004).
DR EMBL; AY600651; AAU11610.1; --
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 12976 MW; 3B44B44EF7601358 CRC64;

Query Match 67.3%; Score 577; DB 2; Length 125;
Best Local Similarity 91.1%; Pred. No. 3.5e-38;
Matches 113; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 1 ARILNLKKKTNTVPTVAHRTSSSRVAVSLVEFTCCRAGALDWVCARRRLPSGRNLEVD 60

Db 2 ARILNLKKKTNTVPTVAHRTSSSRVAVSLVEFTCCRAGALDWVCARRRLPSGRNLEVD 61

QY 61 VSLSLRLVCPRAGGLSGTLGSPAMRAGRGDSCLPVALGLAGAPOTPGVGRAIWR 120

Db 62 VSLSPRRVDPVPGGLSGTLGSPAMRAGRGDSCLPVALGLAGAPOTPGVGRAIWR 121

QY 121 SSIP 124

Db 122 SSIP 125

RESULT 2

Q66VM9 ID Q66VM9 PRELIMINARY; PRT; 125 AA.
AC Q66VM9;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;

Q62RW4 homo sapien
Q91314 aeromonas h
O84HK8 streptomyce
O69949 streptomyce
O6PE19 brachydanio
Q00196 mus musculu
Q987W2 oryza sativ
Q8WY58 homo sapien
Q6NWZ8 homo sapien
Q988U4 rhizobium l
Q8R554 mus musculu
Q9YF32 aeropyrum p
Q64OK4 xenopus lae
Q7Z7Q5 homo sapien

ALIGNMENTS

32 99 11.6 468 2 Q62RW4
33 98.5 11.5 356 2 Q91314
34 97.5 11.4 1960 2 O84HK8
35 97 11.3 363 2 O69949
36 97 11.3 1449 2 O6PE19
37 96 11.2 462 1 PO22 MOUSE
38 95 11.1 527 2 Q987W2
39 94.5 11.0 263 2 Q8WY58
40 94.5 11.0 263 2 Q6NWZ8
41 94.5 11.0 617 2 Q988U4
42 94 11.0 926 1 CE22 MOUSE
43 93.5 10.9 105 2 Q9YF32
44 93 10.9 597 2 Q64OK4
45 92.5 10.8 350 2 Q7Z7Q5


```

[1]
RN SEQUENCE FROM N.A.
RP PubMed=15302945; DOI=10.1099/vir.0.80071-0;
RA Deforges S., Evlashev A., Perret M., Sodoyer M., Pouzol S.,
RA Scoazec J.Y., Bonnaud B., Diaz O., Paranhos-Baccala G., Lotteau V.,
RA Andre P.;
RT "Expression of hepatitis C virus proteins in epithelial intestinal
RT cells in vivo.";
RL J. Gen. Virol. 85:2515-2523(2004).
DR EMBL; AY600671; AAU11630.1; -.
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 13013 MW; D1605136F99F7ECA CRC64;

Query Match 65.0%; Score 557; DB 2; Length 125;
Best Local Similarity 88.7%; Pred. No. 1.3e-36;
Matches 110; Conservative 2; Mismatches 12; Indels 0; Gaps 0;

Qy 1 ARILNKKNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSSGRNLEVD 60
Db 2 SRILNKKENVTPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSSGRNLEVD 61
Qy 61 VSLSPRLVGRPRGCLSPGTLGSPSMARAAAGRGDSCLPVALGLAGAPQTCVGRVIRSSIP 120
Db 62 ASLSPRRVDPRAGCLSLGTLGSPSMARAAAGRGDSCPPAALGLVGLLTPGAGRAIWR 121
Qy 121 SSIP 124
Db 122 SSIP 125

RESULT 3
Q66VK1 PRELIMINARY; PRT; 125 AA.
AC Q66VK1
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15302945; DOI=10.1099/vir.0.80071-0;
RA Deforges S., Evlashev A., Perret M., Sodoyer M., Pouzol S.,
RA Scoazec J.Y., Bonnaud B., Diaz O., Paranhos-Baccala G., Lotteau V.,
RA Andre P.;
RT "Expression of hepatitis C virus proteins in epithelial intestinal
RT cells in vivo.";
RL J. Gen. Virol. 85:2515-2523(2004).
DR EMBL; AY600671; AAU11630.1; -.
FT NON_TER 125 125
SQ SEQUENCE 125 AA; 12951 MW; C66F37EB8C4D7ACA CRC64;

Query Match 64.9%; Score 556; DB 2; Length 125;
Best Local Similarity 88.7%; Pred. No. 1.6e-36;
Matches 110; Conservative 1; Mismatches 13; Indels 0; Gaps 0;

Qy 1 ARILNKKNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSSGRNLEVD 60
Db 2 ARILNKKENVTPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSSGRNLEVD 61
Qy 61 VSLSPRLVGRPRGCLSPGTLGSPSMARAAAGRGDSCLPVALGLAGAPQTCVGRVIRSSIP 120
Db 62 ASLSPRRVDPRAGCLSLGTLGSPSMARAAAGRGDSCPPAALGLVGLLTPGAGRAIWR 121
Qy 121 SSIP 124
Db 122 SSIP 125

RESULT 4
Q66VN1 PRELIMINARY; PRT; 126 AA.
AC Q66VN1
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15302945; DOI=10.1099/vir.0.80071-0;
RA Deforges S., Evlashev A., Perret M., Sodoyer M., Pouzol S.,
RA Scoazec J.Y., Bonnaud B., Diaz O., Paranhos-Baccala G., Lotteau V.,
RA Andre P.;
RT "Expression of hepatitis C virus proteins in epithelial intestinal
RT cells in vivo.";
RL J. Gen. Virol. 85:2515-2523(2004).
DR EMBL; AY600670; AAU11629.1; -.
FT NON_TER 126 126
SQ SEQUENCE 126 AA; 13397 MW; 2A1F8FE1B4A19851 CRC64;

Query Match 50.5%; Score 432.5; DB 2; Length 126;
Best Local Similarity 76.3%; Pred. No. 9.2e-27;
Matches 90; Conservative 2; Mismatches 23; Indels 3; Gaps 2;

Qy 7 KKKTNVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRRLPSSGRNLEVDVLSLSPR 66
Db 12 KRNTNRPPQ-ARQVPGWR--SDQLVEFTCCRAGALDWVCARRRLPSSGRNLEVDASLSPR 68
Qy 67 LVGPBAGCLSPGTLGSPSMARAAAGRGDSCLPVALGLAGAPQTCVGRVIRSSIP 124
Db 69 RVDPRAGCLSLGTLGSPSMARAAAGRGDSCPPAALGLVGLLTPGAGRAIWRSSIP 126

RESULT 5
Q00687 PRELIMINARY; PRT; 122 AA.
AC Q00687
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Core (Fragment).
CN Name-core;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92268871; PubMed=1316939;
RA Chan S., McOmish F., Holmes E., Dow B., Peutherer J., Pollett E.,
RA Yap P., Simmonds P.;
RT "Analysis of a new hepatitis C virus type and its phylogenetic
RT relationship to existing variants.";
RL J. Gen. Virol. 73:1131-1141(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94270990; PubMed=7545932;
RA Hotta H., Doi H., Hayaishi T., Purwanta M., Soemarto W., Mizokami M.,
RA Ohba K., Homma M.;
RT "Analysis of the core and E1 envelope region sequences of a novel
RT variant of hepatitis C virus obtained in Indonesia.";
RL Arch. Virol. 136:53-62(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX Chan S.-W.;
RL Submitted (DEC-1991) to the EMBL/GenBank/DBJ databases.
DR EMBL; D10123; BAA01000.1; -.
FT NON_TER 1 1
```

```
FT NON_TER 122 122
SQ SEQUENCE 122 AA; 12632 MW; PB5C7A7F76AE29AC CRC64;

Query Match
  47.0%; Score 403; DB 2; Length 122;
Best Local Similarity 68.6%; Pred. No. 1.9e-24;
Matches 83; Conservative 9; Mismatches 29; Indels 0; Gaps 0;

Qy 6 LKKTNTVPTVAHRTSSRAVRSVLEFTCCAGALDWVCARRRLPSGRNLEVDVSLSP 65
Db 1 LKEPKETPSVAHRTSSRAVRSVLEFTCCAGALDWVCARRVKKLNGHSLADDDSLSP 60

Qy 66 RLVGPRAGPLSPGTGLGFSMAARAAGRDGSCLPVALGLAGAPOTPGVGRAIWRSSIP 125
Db 61 RRVGAKAGPLSPGTGRGSMVTRAAGGGGSCPHAPVHLGAQMTPGEGPAIWKSSIPS 120

Qy 126 R 126
Db 121 R 121

RESULT 6
Q68874
ID Q68874 PRELIMINARY; PRT; 95 AA.
AC - Q68874;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95270681; PubMed=7751366;
RA Feucht H.H., Zoellner B., Polywka S., Laufs R.;
RT "Study on reliability of commercially available hepatitis C virus
  antibody tests.";
RL J. Clin. Microbiol. 33:620-624(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Feucht H.;
RL Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; X78950; CAA55547.1; -.
DR PIR; S44213; S44213.
FT NON_TER 1 1
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9702 MW; 647C80587C6F992F CRC64;

Query Match
  43.6%; Score 374; DB 2; Length 95;
Best Local Similarity 78.7%; Pred. No. 2.9e-22;
Matches 74; Conservative 2; Mismatches 18; Indels 0; Gaps 0;

Qy 11 NVPTVAHRTSSRAVRSVLEFTCCAGALDWVCARRRLPSGRNLEVDVSLSPRLVGP 70
Db 2 NVPTAAHRTSSRAVRSVLEFTCCAGALDWVCARRRLPSGRNLEVDVSLSPRLASP 61

Qy 71 RAQPGLSPTGLGFSMAARAAGRDGSCLPVALGL 104
Db 62 RAQPGLSPTGLGFSMAARAAGRDGSCLPVALGL 95

RESULT 7
Q70GD2
ID Q70GD2 PRELIMINARY; PRT; 134 AA.
AC Q70GD2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus type 1.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.

OX NCBI_TaxID=41856;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14718617; DOI=10.1099/vir.0.19472-0;
RA Colina R., Casane D., Vasquez S., Garcia L., Chunga A., Romero H.,
  Kahn B., Cristina J.;
RT "Genetic analysis of Hepatitis C viruses in Peruvian patients.";
RL J. Gen. Virol. 85:31-37(2004).
DR EMBL; AJ582128; CAE46584.1; -.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR002522; HCV capsid.
DR Pfam; PF01543; HCV capsid; 1.
FT NON_TER 134 134
SQ SEQUENCE 134 AA; 14214 MW; 85D9075FB32CABAC CRC64;

Query Match
  40.9%; Score 350.5; DB 2; Length 134;
Best Local Similarity 85.4%; Pred. No. 2.9e-20;
Matches 70; Conservative 1; Mismatches 6; Indels 5; Gaps 1;

Qy 52 PSGRNLEVDVSLSPRLVGPAGPLSPGTGLGFSMAARAAGRDGSCLPVALGLAGAPOT 111
Db 58 PRGR-----QTPRHVGPAGPLSPGTGLGFSMAARAAGRDGSCLPVALGLAGAPOT 112

Qy 112 GVGRAIWRSSIPLRAASPTSW 133
Db 113 GVGRAIWRSSIPLRAASPTSW 134

RESULT 8
Q68361
ID Q68361 PRELIMINARY; PRT; 102 AA.
AC Q68361;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=EUUK3;
RX MEDLINE=96030859; PubMed=7595353;
RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
  The Int'l HCV Collaborative Study Group;
RT "Investigation of the pattern of hepatitis C virus sequence diversity
  in different geographical regions: implications for virus
  classification.";
RL J. Gen. Virol. 76:2493-2507(1995).
DR EMBL; U31257; AAA75061.1; -.
FT NON_TER 1 1
FT NON_TER 102 102
SQ SEQUENCE 102 AA; 10446 MW; D3C38355326B429D CRC64;

Query Match
  40.4%; Score 346; DB 2; Length 102;
Best Local Similarity 69.9%; Pred. No. 5.1e-20;
Matches 72; Conservative 6; Mismatches 23; Indels 2; Gaps 2;

Qy 9 KTNVTPTVAHRTSSRAVRSVLEFTCCAGALDWVCARRRLPSGRNLEVDVSLSPRLV 68
Db 1 KTNVHTAAHRTSSRAVRSVLEFTCCAGALDWVCARRRLPSGRNLEVDVSLSPRLV 60

Qy 69 GPRAGPLSPGTGLGFSMAARAAGRDGSCLPV-ALGLAGAPOT 110
Db 61 SPRAGGVNPGTLPPTMRASGGQ-GGCSPLEALGLIGAPMT 102

RESULT 9
Q68358
ID Q68358 PRELIMINARY; PRT; 102 AA.
AC Q68358;
```

```

DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EUBAH3485;
RX MEDLINE=96030859; PubMed=7595353;
RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA The Int'l HCV Collaborative Study Group;
RT "Investigation of the pattern of hepatitis C virus sequence diversity
RT in different geographical regions: implications for virus
RT classification.";
RL J. Gen. Virol. 76:2493-2507(1995).
DR EMBL; U31254; AAA75058.1; -.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 10407 MW; E3B0611FF5880B35 CRC64;

Query Match 39.2%; Score 336; DB 2; Length 102;
Best Local Similarity 69.7%; Pred. No. 3.1e-19;
Matches 69; Conservative 2; Mismatches 28; Indels 0; Gaps 0;

Qy 9 KTNVTPTVAHRTSSRVAVRSIVFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68
Db 1 KPNVTPTAAPWTSSRAVRSIAEFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68

Qy 69 GPRAGPGLSPGTLGPGSMAMRAAGRGDSCLPVALGLAGA 107
Db 61 GPRAGPGLSLGILGFTTMRVAGGGGSCHPAVALGRLGA 99

RESULT 10
Q68356 PRELIMINARY; PRT; 102 AA.
AC Q68356;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EUBAH3485;
RX MEDLINE=96030859; PubMed=7595353;
RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA The Int'l HCV Collaborative Study Group;
RT "Investigation of the pattern of hepatitis C virus sequence diversity
RT in different geographical regions: implications for virus
RT classification.";
RL J. Gen. Virol. 76:2493-2507(1995).
DR EMBL; U31252; AAA75056.1; -.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 10539 MW; 308EC3CDB0F80E96 CRC64;

Query Match 38.7%; Score 331.5; DB 2; Length 102;
Best Local Similarity 68.4%; Pred. No. 7.1e-19;
Matches 67; Conservative 7; Mismatches 21; Indels 3; Gaps 1;

Qy 9 KTNVTPTVAHRTSSRVAVRSIVFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68
Db 1 KPNVTPVAHRTSSRVAVRSIVFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68

Qy 69 GPRAGPGLSPGTLGPGSMAMRAAGRGDSC---LPVALG 103
Db 1 KPNVTPVAHRTSSRVAVRSIVFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68

Query Match 35.6%; Score 305; DB 2; Length 100;
Best Local Similarity 67.0%; Pred. No. 8.6e-17;
Matches 65; Conservative 2; Mismatches 30; Indels 0; Gaps 0;
```

```

Db 61 GAKAGPGLSPGTLGPGSMVMTTRAAGGQSGSCPHAAPVQLG 98

RESULT 11
Q68359 PRELIMINARY; PRT; 102 AA.
AC Q68359;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EUEG15;
RX MEDLINE=96030859; PubMed=7595353;
RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA The Int'l HCV Collaborative Study Group;
RT "Investigation of the pattern of hepatitis C virus sequence diversity
RT in different geographical regions: implications for virus
RT classification.";
RL J. Gen. Virol. 76:2493-2507(1995).
DR EMBL; U31255; AAA75059.1; -.
FT NON_TER 1
FT NON_TER 102
SQ SEQUENCE 102 AA; 10607 MW; C6F24F1A392EEB87 CRC64;

Query Match 37.2%; Score 319; DB 2; Length 102;
Best Local Similarity 68.4%; Pred. No. 6.9e-18;
Matches 67; Conservative 2; Mismatches 29; Indels 0; Gaps 0;

Qy 9 KTNVTPTVAHRTSSRVAVRSIVFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68
Db 1 KPNVTPTAAPWTSSRAVRSIAEFTCCRAGALDWCARRRLPSGRNLEVDVLSPLRV 68

Qy 69 GPRAGPGLSPGTLGPGSMAMRAAGRGDSCLPVALGLAG 106
Db 61 GPRAGPGLSLGILGFTTMRVAGGGGSCHPAALGRPG 98

RESULT 12
Q68360 PRELIMINARY; PRT; 100 AA.
AC Q68360;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Core protein (Fragment).
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EUEG28;
RX MEDLINE=96030859; PubMed=7595353;
RA Mellor J., Holmes E.C., Jarvis L.M., Yap P.L., Simmonds P.,
RA The Int'l HCV Collaborative Study Group;
RT "Investigation of the pattern of hepatitis C virus sequence diversity
RT in different geographical regions: implications for virus
RT classification.";
RL J. Gen. Virol. 76:2493-2507(1995).
DR EMBL; U31256; AAA75060.1; -.
FT NON_TER 1
FT NON_TER 100
SQ SEQUENCE 100 AA; 10414 MW; 727D8920E9BE281C CRC64;

Query Match 35.6%; Score 305; DB 2; Length 100;
Best Local Similarity 67.0%; Pred. No. 8.6e-17;
Matches 65; Conservative 2; Mismatches 30; Indels 0; Gaps 0;
```

Search completed: August 25, 2005, 03:36:22
Job time : 173 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 02:39:56 ; Search time 22 Seconds
(without alignments)
556.475 Million cell updates/sec

Title: US-10-664-038-1

Perfect score: 857

Sequence:

1 ARILNLKKTNTPTVAHRT.....ALPGWRMASGFWKTATMQQ 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pap.*
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pap.*
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pap.*
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pap.*
- 5: /cgn2_6/prodata/1/iaa/6CTUS COMB.pap.*
- 6: /cgn2_6/prodata/1/iaa/backfile1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	100.0	164	4	US-09-644-987-1
2	334	37.8	115	4	US-09-878-281A-148
3	100	11.7	236	4	US-09-252-991A-27618
4	99.5	11.6	263	4	US-09-252-991A-16906
5	98	11.4	162	4	US-09-252-991A-18207
6	98	11.4	202	4	US-09-252-991A-24790
7	97	11.3	231	4	US-09-252-991A-22801
8	96	11.2	709	4	US-09-252-991A-27305
9	95.5	11.1	328	4	US-09-252-991A-17729
10	95	11.1	371	4	US-09-252-991A-25006
11	93	10.9	213	4	US-09-252-991A-20428
12	92.5	10.8	254	4	US-09-252-991A-23311
13	91.5	10.7	461	4	US-09-252-991A-24717
14	91	10.6	253	4	US-09-252-991A-25632
15	91	10.6	1027	4	US-09-252-991A-26216
16	90	10.5	220	4	US-09-252-991A-16739
17	90	10.5	242	4	US-09-252-991A-31425
18	89.5	10.4	138	4	US-09-252-991A-26931
19	89.5	10.4	1209	4	US-09-252-991A-25844
20	89.5	10.4	1225	4	US-09-252-991A-25018
21	89	10.4	348	4	US-09-252-991A-25605
22	89	10.4	425	4	US-09-252-991A-20467
23	89	10.4	470	4	US-09-949-016-9859
24	88.5	10.3	171	4	US-09-252-991A-23216
25	88	10.3	146	4	US-09-252-991A-27352
26	88	10.3	504	4	US-09-252-991A-28242
27	87.5	10.2	205	4	US-09-252-991A-17563

ALIGNMENTS

RESULT 1

US-09-644-987-1
; Sequence 1, Application US/09644987
; Patent No. 6803214
; GENERAL INFORMATION:
; APPLICANT: MAVROMARA, PENELOPE
; APPLICANT: VARAKLIOTI, AGORITSA
; APPLICANT: GEORGIOPOULOU, URANIA
; TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH
; TITLE OF INVENTION: AND/OR OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: 03495-0194-00000
; CURRENT APPLICATION NUMBER: US/09/644,987
; CURRENT FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/151,074
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-644-987-1

Query Match	100.0%;	Score	857;	DB	4;	Length	164;
Best Local Similarity	100.0%;	Pred. No.	5.8e-82;				
Mismatches	164;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
QY	1	ARILNLKKTNTPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWCARRERLPSSGRNLEVD	60				
DB	1	ARILNLKKTNTPTVAHRTSSSRVAVRSLSVEFTCCRAGALDWCARRERLPSSGRNLEVD	60				
QY	61	VSLSPRLVGRAGPGLSPGTLGPSNMAAAGRGDSCILPVALGLAGAPQTPGVGRAIWR	120				
DB	61	VSLSPRLVGRAGPGLSPGTLGPSNMAAAGRGDSCILPVALGLAGAPQTPGVGRAIWR	120				
QY	121	SSIPLRAASPTSWGTYRSSAPLLEALPGWRMASGFWKTATMQQ	164				
DB	121	SSIPLRAASPTSWGTYRSSAPLLEALPGWRMASGFWKTATMQQ	164				

RESULT 2

US-09-878-281A-148
; Sequence 148, Application US/09878281A
; Patent No. 6762024
; GENERAL INFORMATION:
; APPLICANT: Imogenetics N.V.
; TITLE OF INVENTION: New sequences of hepatitis C virus genotypes for diagnosis, prophylaxis and therapy
; TITLE OF INVENTION: and therapy
; FILE REFERENCE: 35

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24790

; LENGTH: 202
; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24790

Query Match 11.4%; Score 98; DB 4; Length 202;
Best Local Similarity 25.1%; Pred. No. 0.02;

Matches 46; Conservative 12; Mismatches 71; Indels 54; Gaps 7;
Qy 14 PTVAHRTSSSRVAVRSVLEFTCCRAGALDWCVA-----RRRLPGRNLEVDVLSLSPR 66

Db 1 PSTSWASPSRLPVRP-----CSTCMPWACAGSRATTWAPPAAPSAAWATAAWSPS 54
Qy 67 LVDPAGPGL-----SPG-----TLGPSMAMRAAGRDGSC 98

Db 55 TAGPSAAPAPSPSSGRACGWKPKATASTRRNGHEPAPGDRRRRLGRHGRSHRAGQARGPLR 114
Qy 99 PVALGLA---CAPQTPGVGRAIWRSSITPLRAASPTSGT-----YRSSAPLLEALPG 148

Db 115 PLRGLASRRGGLSRPLAGRR---RSGLFRRLHPDAGKTAARFLRLRRHRAPEQPRG 171
Qy 149 PWR 151
Db 172 RWR 174

RESULT 7

US-09-252-991A-22801

; Sequence 22801, Application US/09252991A
; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 22801

; LENGTH: 231

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-22801

Query Match 11.3%; Score 97; DB 4; Length 231;
Best Local Similarity 29.6%; Pred. No. 0.03;

Matches 45; Conservative 8; Mismatches 55; Indels 44; Gaps 7;
Qy 10 TNYTPTVAHRTSSSRVAVRSVLEFTCCRAGALDW--VCARRERLPGRNLEVDVLSLSPRL 67

Db 8 TRTAGACSSRCSSAR-----CHSACATRWFGCAARSALPSPGR-----RG 47
Qy 68 VGPRAGPGLSPGTLGPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWRASSIPLRA 127

Db 48 SGTAGTNCWFSPTATSMPTACLARST---TAANIRAAMP-----PLTV 91
Qy 128 ASPTS-----NGTYRSSAPLLEALPGPWRMASG 155

Db 92 SSPTSSRRPWPH--SMMWREAIRAWWRIASG 121

RESULT 8

US-09-252-991A-27305

; Sequence 27305, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 27305

; LENGTH: 709

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-27305

Query Match 11.2%; Score 96; DB 4; Length 709;
Best Local Similarity 34.5%; Pred. No. 0.16;

Matches 50; Conservative 6; Mismatches 65; Indels 24; Gaps 9;
Qy 19 RTSSSRVAVRSVLEFTCCRAGALDWVCARRERLPGRNLEVDVLSLSP-----RLVGPR 71

Db 96 RTFGGRPACKRVA---TRARRSAL--ADAGRGGLPAGRRORPDALARPAPRGAGERLSRSR 150
Qy 72 AGPGLSPGTLGPSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWRSSIPLRAASPT 131

Db 151 AVAALRPSS--RPALARRAGGAGGA--PPEPGVALEPDPGAGR--WLPGG-AFPAAAPG 204
Qy 132 SWGTYRSSAPLLEALPGPWRMASGF 156

Db 205 RLG----SRP--GAQPGPGLRAGGF 223

RESULT 9

US-09-252-991A-17729

; Sequence 17729, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 17729

; LENGTH: 328

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17729

Query Match 11.1%; Score 95.5; DB 4; Length 328;
Best Local Similarity 29.4%; Pred. No. 0.067;

Matches 48; Conservative 12; Mismatches 60; Indels 43; Gaps 8;
Qy 14 PTVAHRTSSSRVAVRSVLEFTCCRAGALDWVCARRERLPSPGR-----NLEVDV 61

Db 95 PPVAHRPSRR-----PHGA-----AR-----PAGRRSGPGQDHRGPDHPSPV 133
Qy 62 SLSPRLVGPRAGPGLSPGTL--GPSMAMRAAGG--RDGS-----CLPVALGLAGAPQTPG 112

Db 134 ALRERRAGDPCKPAPVVGDPAPALQAGGVPVQCALRRKRCQQLRGHPVGPGRPG 193
 QY 113 VGRAIWRSSPLRAASPTSGTYESSAPLLEALPGPWRMASG 155
 Db 194 VAEGRRTRPGPVRRLRGPAGGR-RSPPPGLAPGFGORIRAG 235

RESULT 10
 US-09-252-991A-25006
 ; Sequence 25006, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 25006
 ; LENGTH: 371
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-25006

Query Match 11.1%; Score 95; DB 4; Length 371;
 Best Local Similarity 32.6%; Pred. No. 0.089;
 Matches 44; Conservative 10; Mismatches 53; Indels 28; Gaps 6;

QY 47 RRRLPSGRNLEVDVSLSPRLVGPAGGLSPGLGSPMSAMRAGGRDGLPVALGLAG 106
 Db 31 RRDPGGRVRLRQVADRPAPAPAGRG-APG-----RQAPGRGTARPVAAAG 82

QY 107 ---APQTPGVGRATWRSSI-----PLR---AASPTSW-----GTYRSSAPLLEAL 146
 Db 83 DSWAHRHGVGADDLAQSIAMHRPPGRRGPAPASPAWRRAARAGTGARTGRPAAAT 142

QY 147 GGPWRMASGFWKTAT 161
 Db 143 PARLSAPALRRPAT 157

RESULT 11
 US-09-252-991A-24258
 ; Sequence 24258, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; PRIOR FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 24258
 ; LENGTH: 213
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-24258

Query Match 10.9%; Score 93; DB 4; Length 213;
 Best Local Similarity 29.8%; Pred. No. 0.07;
 Matches 54; Conservative 13; Mismatches 60; Indels 54; Gaps 10;

QY 15 TVAHRTSSSRV--AVRSLVEFTCCR-----AGALDWCARRERLPSGRNLEVDVLSLSPR 66
 Db 34 TPARRSSGRRLPALRLHTQRPVAVRHQGPQRPAGTL-----RRERLLRIRRR-----AHRRR 84

QY 67 LVGPFRAGGLSPGLTGLSPMSAMRAG--GRDGLPVALGLAGAPQTPGV-----GRAIWR 120
 Db 85 LVGDROGGRRAGSGDAAVVLRQRQTAGRPSTRYPTAAC-A-PFSAPVLRRLRRPGRAGARR 143

QY 121 SSIPLRA--ASPTSW-----GTYRSSA-----PILLEALPGPW 150
 Db 144 GSPFGSPRRSPGSPWPCSPACRCGRPCPCPGCHGTAPPAAGSLPRPPAPWRSPPAPRLPGPW 203

QY 151 R 151
 Db 204 R 204

RESULT 12
 US-09-252-991A-23311
 ; Sequence 23311, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 23311
 ; LENGTH: 254
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-23311

Query Match 10.8%; Score 92.5; DB 4; Length 254;
 Best Local Similarity 29.6%; Pred. No. 0.099;
 Matches 50; Conservative 12; Mismatches 80; Indels 27; Gaps 8;

QY 13 TPTVAHRTSSSRVAVRSLVEFTCC-----RAGALDWCARRERLPS-GRNLEVDV 61
 Db 86 TPAPARAGRADRLVRGPALRSALPGSAARQHARRQSGATLQCARPRQPAQGRRRHPS 145

QY 62 SLSP-----RLVGPFRAGGLSPGLTGLSPMSAMRAGGRDGLPVALGLAGAP-QTPGV 114
 Db 146 ATAPGTAPRFRRLGPRSPRAAAHWPAPVAGHPGRPEPLLPGAGRHRHPRVPAQGVG 205

QY 115 RAIWVRSSTP-----LRAASP-TSMGTYSRSPAPLLEALPG-PWRMASG 155
 Db 206 -ATGLRADCPAGVYTLRRGRPEPAGCAARGTGGLRLPGSPFAETARG 253

RESULT 13
 US-09-252-991A-24717
 ; Sequence 24717, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 24717


```
; LENGTH: 461 ;
; TYPE: PRT ;
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24717

Query Match      10.7%; Score 91.5; DB 4; Length 461;
Best Local Similarity 28.2%; Pred. No. 0.27;
Matches 42; Conservative 14; Mismatches 54; Indels 39; Gaps 7;

QY 44 VCARRERLPSGR-----NLEVDVLSR-----LVG-----PRAGP-----GL--- 76
Db 228 VAGRRRLPGKLAGAALRRRLRRRCRGLAGNGLATVPAGPPGQRVGLPAC 287
QY 77 -----SPGTGFSMAMRAAGRDGSCLPVALGLAGAPQTPGVGRAIWRSSIPLR 126
Db 288 LRVTRDRLRPPVAALPEVARSGFAPS GTA-PAGGGAAGKSPAGLGR-LARRSSFPSS 345
QY 127 AASPTSGTYRSSAPLLEALPGPWRMASG 155
Db 346 SATPPAATTYFAARAPMPAITEPGRS GTG 374

RESULT 14
US-09-252-991A-29632
; Sequence 29632, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29632
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29632

Query Match      10.6%; Score 91; DB 4; Length 253;
Best Local Similarity 32.6%; Pred. No. 0.14;
Matches 43; Conservative 10; Mismatches 41; Indels 38; Gaps 8;

QY 46 ARHERLPGRNLEVD-----VLSPLRVLPGRAGPGLSPGTGLGPS-----MAMR 88
Db 131 ARSPGPAGRAVRPGAAGAGATGPPAVRGRGVRVREGAGCEVAPGATRSSSSAPGRAGR 190
QY 89 AAGRGDGSCLP-----VALGLAGAPQT-PGVGRAIWRSSIPLRAS-PTSW-----GTYS 138
Db 191 PAGRREGSGRPGRRERRGRSSAPRSRGTTGRA-----CRAGRGTSWRRLGTGRC 242
QY 139 SAPLLEALPGPW 150
Db 243 RSP---ASPGGW 251

RESULT 15
US-09-252-991A-26216
; Sequence 26216, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
```

```
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26216
; LENGTH: 1027
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26216

Query Match      10.6%; Score 91; DB 4; Length 1027;
Best Local Similarity 28.9%; Pred. No. 0.86;
Matches 48; Conservative 8; Mismatches 56; Indels 54; Gaps 8;

QY 36 CRAGALDWVCARRERL-----PSGRNLEVDVLSLPRLVG----- 69
Db 179 CRGNAV P--PDRATRLAAAGAPAGGE-----SLSPRAGGLELGRPAPAGAGLPAAARR 231
QY 70 --PRAGPGLSPGTGLGFSMAMRAAGR--DGSCLPVALGLAGAPQTPGV-----G 114
Db 232 AIPRACRG--GCRGPGARAGKHGRRLPGTARLPAGRLAAAPALSGVFPWPGAGEGLR 288
QY 115 RAIWRSSIPLRASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 160
Db 289 RARWGRPVAGIRAPRR---GSRKSPGVLVPCSPGTLRAAAALWPGA 331

Search completed: August 25, 2005, 03:18:11
Job time : 23 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 25, 2005, 03:08:37 ; Search time 161 Seconds

(without alignments)
398.883 Million cell updates/sec

Title: US-10-664-038-1

Perfect score: 857

Sequence: 1 ARILNLKKTNTVPTVAHRT.....ALPGPWRMASGFWKATMQQ 164

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1759131 seqs, 391586102 residues

Total number of hits satisfying chosen parameters: 1759131

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US05_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	857	100.0	164	17	US-10-664-038-1
2	837	97.7	160	15	US-10-150-283-1
3	824	96.1	196	16	US-10-601-020-2
4	824	96.1	196	16	US-10-601-020-10
5	806	94.0	161	9	US-09-736-959A-1
6	806	94.0	161	9	US-09-736-959A-32
7	768	89.6	161	9	US-09-736-959A-3
8	763	89.0	161	9	US-09-736-959A-2
9	658	76.8	192	16	US-10-601-020-17
10	648.5	75.7	195	16	US-10-601-020-9
11	624	72.8	161	9	US-09-736-959A-7

12	583.5	68.1	195	16	US-10-601-020-13	Sequence 13, Appl
13	552	64.4	143	9	US-09-736-959A-5	Sequence 5, Appl
14	548	63.9	143	9	US-09-736-959A-9	Sequence 9, Appl
15	547	63.8	143	9	US-09-736-959A-8	Sequence 8, Appl
16	540	63.0	143	9	US-09-736-959A-4	Sequence 4, Appl
17	540	63.0	143	9	US-09-736-959A-11	Sequence 10, Appl
18	540	63.0	143	9	US-09-736-959A-11	Sequence 11, Appl
19	534	62.3	143	9	US-09-736-959A-12	Sequence 12, Appl
20	490	57.2	195	16	US-10-601-020-16	Sequence 16, Appl
21	487	56.8	154	9	US-09-736-959A-16	Sequence 16, Appl
22	480.5	56.1	192	16	US-10-601-020-11	Sequence 11, Appl
23	475	55.4	194	16	US-10-601-020-15	Sequence 15, Appl
24	444	51.8	197	16	US-10-601-020-12	Sequence 12, Appl
25	428	49.9	139	9	US-09-736-959A-6	Sequence 6, Appl
26	424	49.5	154	9	US-09-736-959A-15	Sequence 15, Appl
27	417	48.7	139	9	US-09-736-959A-17	Sequence 17, Appl
28	409	47.7	125	9	US-09-736-959A-20	Sequence 20, Appl
29	384.5	44.9	194	16	US-10-601-020-14	Sequence 14, Appl
30	365	42.6	125	9	US-09-736-959A-19	Sequence 19, Appl
31	363	42.4	125	9	US-09-736-959A-21	Sequence 21, Appl
32	353	41.2	125	9	US-09-736-959A-14	Sequence 14, Appl
33	339	39.6	125	9	US-09-736-959A-18	Sequence 18, Appl
34	324	37.8	115	10	US-09-873-224-148	Sequence 148, App
35	306	35.7	125	9	US-09-736-959A-13	Sequence 13, Appl
36	113	13.2	19608	15	US-10-084-846A-8	Sequence 8, Appl
37	95	11.1	527	16	US-10-437-963-114138	Sequence 114138,
38	95	11.1	19695	15	US-10-084-846A-3	Sequence 3, Appl
39	93.5	10.9	151	15	US-10-108-260A-4036	Sequence 4036, Ap
40	93.5	10.9	602	15	US-10-104-047-3324	Sequence 3324, Ap
41	92.5	10.8	705	14	US-10-156-761-13190	Sequence 13190, A
42	90.5	10.6	179	16	US-10-425-115-251407	Sequence 251407,
43	90.5	10.6	205	16	US-10-437-963-114227	Sequence 114227,
44	90.5	10.6	291	15	US-10-094-749-2648	Sequence 2648, Ap
45	90	10.5	217	15	US-10-425-114-46623	Sequence 46623, A

ALIGNMENTS

RESULT 1

US-10-664-038-1
; Sequence 1, Application US/10864038
; Publication No. US20050053915A1
; GENERAL INFORMATION:
; APPLICANT: NAVROMARA, PENELOPE
; APPLICANT: VARAKLIOTI, AGORITSA
; APPLICANT: GEORGIOPOULOU, URANIA
; TITLE OF INVENTION: NUCLEIC ACIDS AND NEW POLYPEPTIDES ASSOCIATED WITH
; TITLE OF INVENTION: AND/OR OVERLAPPING WITH HEPATITIS C VIRUS CORE GENE
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: 03495-0194-00000
; CURRENT APPLICATION NUMBER: US/10/664,038
; CURRENT FILING DATE: 2003-07-17
; PRIOR APPLICATION NUMBER: US/09/644,987
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: 60/151,074
; PRIOR FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-664-038-1

Query Match 100.0%; Score 857; DB 17; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.7e-70;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARILNLKKTNTVPTVAHRTSSRVAVSLVFTCCRAGALDWCARRLPSPGRNLEVD 60

DB 1 ARILNLKKTNTVPTVAHRTSSRVAVSLVFTCCRAGALDWCARRLPSPGRNLEVD 60

QY 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
Db 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164
Db 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164

RESULT 2

US-10-150-283-1
; Sequence 1, Application US/10150283
; Publication No. US20030219407A1
; GENERAL INFORMATION:
; APPLICANT: Ding, Shou-wei
; APPLICANT: Li, Hong-wei
; APPLICANT: Li, Wan-xiang
; TITLE OF INVENTION: The Regents of the University of California
; FILE REFERENCE: RNA Silencing in Animals as an Antiviral Defense
; CURRENT APPLICATION NUMBER: US/10/150,283
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 160
; TYPE: PRT
; ORGANISM: Hepatitis C virus
; FEATURE:
; OTHER INFORMATION: F protein
US-10-150-283-1

Query Match 97.7%; Score 837; DB 15; Length 160;
Best Local Similarity 100.0%; Pred. No. 1.1e-68; Indels 0; Gaps 0;
Matches 160; Conservative 0; Mismatches 0

QY 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
Db 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
QY 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
Db 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 160
Db 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 160

RESULT 3

US-10-601-020-2
; Sequence 2, Application US/10601020
; Publication No. US20040156862A1
; GENERAL INFORMATION:
; APPLICANT: Branch, Andrea D.
; APPLICANT: Walewski, Jose L.
; APPLICANT: Stump, DeCherd D.
; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
; FILE REFERENCE: R11-003CPUSCN
; CURRENT APPLICATION NUMBER: US/10/601,020
; CURRENT FILING DATE: 2003-06-20
; PRIOR FILING DATE: US 09/719277
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/088670
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: US 60/089138
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: PCT/US99/12929
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 196

; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-601-020-2

Query Match 96.1%; Score 824; DB 16; Length 196;
Best Local Similarity 96.3%; Pred. No. 2.2e-67;
Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
Db 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
QY 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
Db 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164
Db 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164

RESULT 4

US-10-601-020-10
; Sequence 10, Application US/10601020
; Publication No. US20040156862A1
; GENERAL INFORMATION:
; APPLICANT: Branch, Andrea D.
; APPLICANT: Walewski, Jose L.
; APPLICANT: Stump, DeCherd D.
; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
; FILE REFERENCE: R11-003CPUSCN
; CURRENT APPLICATION NUMBER: US/10/601,020
; CURRENT FILING DATE: 2003-06-20
; PRIOR FILING DATE: US 09/719277
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: US 60/088670
; PRIOR FILING DATE: 1998-06-09
; PRIOR APPLICATION NUMBER: US 60/089138
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: PCT/US99/12929
; PRIOR FILING DATE: 1999-06-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 10
; LENGTH: 196
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-10-601-020-10

Query Match 96.1%; Score 824; DB 16; Length 196;
Best Local Similarity 96.3%; Pred. No. 2.2e-67;
Matches 158; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
Db 1 ARILNLKKTNTVPTVAHRTSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
QY 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
Db 61 VLSPLRVGPRAGPGLSGTTLGSPNMAAAGRGDSCLPVALGLAGAPQTPGVGRAIWR 120
QY 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164
Db 121 SSIPRAASPTSWGTYRSSAPLLEALPGPWRMASGFWKATMQQ 164

RESULT 5

US-09-736-959A-1
; Sequence 1, Application US/09736959A
; Patent No. US20020076415A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: OU, Jing-Haiung

APPLICANT: XU, Zhenming
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
; FILE REFERENCE: 13761-742
; CURRENT APPLICATION NUMBER: US/09/736,959A
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/170,835
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 161
; TYPE: PR1
; ORGANISM: Hepatitis C Virus
; FEATURE:
; OTHER INFORMATION: Exemplary P17 sequences
US-09-736-959A-1

Query Match 94.0%; Score 806; DB 9; Length 161;
Best Local Similarity 99.4%; Pred. No. 7.7e-66;
Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 66
DB 8 QKTNVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 67
QY 67 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 126
DB 68 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 127
QY 127 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 160
DB 128 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 161

RESULT 6
US-09-736-959A-32
; Sequence 32, Application US/09736959A
; Patent No. US20020076415A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: OU, Jing-Hsiung
; APPLICANT: XU, Zhenming
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
; FILE REFERENCE: 13761-742
; CURRENT APPLICATION NUMBER: US/09/736,959A
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/170,835
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 161
; TYPE: PR1
; ORGANISM: Hepatitis C Virus (HCV-1 subtype)
; FEATURE:
; OTHER INFORMATION: Predicted P17 sequence
US-09-736-959A-32

Query Match 94.0%; Score 806; DB 9; Length 161;
Best Local Similarity 99.4%; Pred. No. 7.7e-66;
Matches 153; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 66
DB 8 QKTNVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 67
QY 67 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 126
DB 68 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 127
QY 127 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 160
DB 128 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 161

RESULT 7
US-09-736-959A-3
; Sequence 3, Application US/09736959A
; Patent No. US20020076415A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: OU, Jing-Hsiung
; APPLICANT: XU, Zhenming
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
; FILE REFERENCE: 13761-742
; CURRENT APPLICATION NUMBER: US/09/736,959A
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/170,835
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 161
; TYPE: PR1
; ORGANISM: Hepatitis C Virus
; FEATURE:
; OTHER INFORMATION: Exemplary P17 sequences
US-09-736-959A-3

Query Match 89.6%; Score 768; DB 9; Length 161;
Best Local Similarity 95.5%; Pred. No. 2.3e-62;
Matches 147; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 66
DB 8 QKTNVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 67
QY 67 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 126
DB 68 LVGPRAGPGLSPGTGPGSMAMRAAGRGDSCPLVALGLAGAPQTPGVGRAIWRSSIPLR 127
QY 127 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 160
DB 128 AASPTSGTGRSSAPLLEALPGPWRMASGFWKTA 161

RESULT 8
US-09-736-959A-2
; Sequence 2, Application US/09736959A
; Patent No. US20020076415A1
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: OU, Jing-Hsiung
; APPLICANT: XU, Zhenming
; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
; FILE REFERENCE: 13761-742
; CURRENT APPLICATION NUMBER: US/09/736,959A
; CURRENT FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/170,835
; PRIOR FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 161
; TYPE: PR1
; ORGANISM: Hepatitis C Virus
; FEATURE:
; OTHER INFORMATION: Exemplary P17 sequences
US-09-736-959A-2

Query Match 89.0%; Score 763; DB 9; Length 161;
Best Local Similarity 94.8%; Pred. No. 6.7e-62;
Matches 146; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 66
DB 8 QKTNVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRERLPGRNLEVDVLSR 67

Db 8 QKQNVPTVAHRTSSGAVVAVRSVLEFTCCRAGALDWCARRGLPSGRNLEVDVLSR 67
 QY 67 LVGPRAGPGLSPGTLGFSMAMRAAGGDSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 126
 Db 68 HVGPRAGPGLSPGTLGFSMAMRVAGGDSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 127
 QY 127 AASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 160
 Db 128 AASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 161

RESULT 9
 US-10-601-020-17
 ; Sequence 17, Application US/10601020
 ; Publication No. US20040156862A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Branch, Andrea D.
 ; APPLICANT: Walewski, Jose L.
 ; APPLICANT: Stump, DeChard D.
 ; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
 ; FILE REFERENCE: RII-003CPUSCN
 ; CURRENT APPLICATION NUMBER: US/10/601,020
 ; CURRENT FILING DATE: 2003-06-20
 ; PRIOR APPLICATION NUMBER: US 09/719277
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: US 60/088670
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: US 60/089138
 ; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US99/12929
 ; PRIOR FILING DATE: 1999-06-09
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 17
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Hepatitis C virus
 US-10-601-020-17

Query Match 76.8%; Score 658; DB 16; Length 192;
 Best Local Similarity 77.4%; Pred. No. 3.3e-52;
 Matches 127; Conservative 9; Mismatches 26; Indels 2; Gaps 1;
 QY 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
 Db 1 AQLNLKEKNVPTAAHRTSSSRVAVRSVLEFTCCRAGAPGWVCARLGLPSGRNLEVD 60
 QY 61 VLSPLRVGPRAGPGLSPGTLGFSMAMRAAGGDSCLPVALGLAGAPOTPGVGRAIWR 120
 Db 61 DNLSPRAGPAGPGLSPGTLGFSMAMRVAGGDSCHPGALGLVGPAPRTPGVGRIWVR 120
 QY 121 SSIPRAASPTSGTYRSSAPLLEALPGPWRMASGFWKATMOQ 164
 Db 121 SSIPRHAGSPTSGTYRSSAP--GALPGPWRMVSGFWRSATTOQ 162

RESULT 10
 US-10-601-020-9
 ; Sequence 9, Application US/10601020
 ; Publication No. US20040156862A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Branch, Andrea D.
 ; APPLICANT: Walewski, Jose L.
 ; APPLICANT: Stump, DeChard D.
 ; TITLE OF INVENTION: NOVEL HEPATITIS C VIRUS PEPTIDES AND USES THEREOF
 ; FILE REFERENCE: RII-003CPUSCN
 ; CURRENT APPLICATION NUMBER: US/10/601,020
 ; CURRENT FILING DATE: 2003-06-20
 ; PRIOR APPLICATION NUMBER: US 09/719277
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: US 60/088670
 ; PRIOR FILING DATE: 1998-06-09
 ; PRIOR APPLICATION NUMBER: US 60/089138

; PRIOR FILING DATE: 1998-06-11
 ; PRIOR APPLICATION NUMBER: PCT/US99/12929
 ; PRIOR FILING DATE: 1999-06-09
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 195
 ; TYPE: PRT
 ; ORGANISM: Hepatitis C virus
 US-10-601-020-9

Query Match 75.7%; Score 648.5; DB 16; Length 195;
 Best Local Similarity 78.7%; Pred. No. 2.5e-51;
 Matches 129; Conservative 6; Mismatches 28; Indels 1; Gaps 1;
 QY 1 ARILNLKKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVD 60
 Db 1 AQLNLKEKNVPTAAHRTSSSRVAVRSVLEFTCCRAGAPGWVCARLGLPSGRNLEVD 60
 QY 61 VLSPLRVGPRAGPGLSPGTLGFSMAMRAAGGDSCLPVALGLAGAPOTPGVGRAIWR 120
 Db 61 ASLSPRIAGPAGPGLSPGTLGFSMAMRVAGGDSCHPGALGLVGPAPRTPGVGRIWVR 120
 QY 121 SSIPRAASPTSGTYRSSAPLLEALPGPWRMASGFWKATMOQ 164
 Db 121 SSIPRAASPTSGTYRSSAP--LGASPELWRMASGWRGTITQQ 163

RESULT 11
 US-09-736-959A-7
 ; Sequence 7, Application US/09736959A
 ; Patent No. US20020076415A1
 ; GENERAL INFORMATION:
 ; APPLICANT: University of Southern California
 ; APPLICANT: XU, Jing-Hsiung
 ; APPLICANT: XU, Zhenming
 ; TITLE OF INVENTION: HEPATITIS C VIRUS GENE PRODUCTS
 ; FILE REFERENCE: 13761-742
 ; CURRENT APPLICATION NUMBER: US/09/736,959A
 ; CURRENT FILING DATE: 2000-12-14
 ; PRIOR APPLICATION NUMBER: US 60/170,835
 ; PRIOR FILING DATE: 1999-12-14
 ; NUMBER OF SEQ ID NOS: 32
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7
 ; LENGTH: 161
 ; TYPE: PRT
 ; ORGANISM: Hepatitis C virus
 ; FEATURE:
 ; OTHER INFORMATION: Exemplary P17 sequences
 US-09-736-959A-7

Query Match 72.8%; Score 624; DB 9; Length 161;
 Best Local Similarity 77.3%; Pred. No. 3.5e-49;
 Matches 119; Conservative 7; Mismatches 28; Indels 0; Gaps 0;
 QY 7 KKTNTVPTVAHRTSSSRVAVRSVLEFTCCRAGALDWCARRRLPSGRNLEVDVLSR 66
 Db 8 QKQNVPTAAHRTSSSRVAVRSVLEFTCCRAGAPGWVCARLGLPSGRNLEVDNLSR 67
 QY 67 LVGPRAGPGLSPGTLGFSMAMRAAGGDSCLPVALGLAGAPOTPGVGRAIWRSSIPLR 126
 Db 68 LADPRAGPGLSPGTLGFSMAMRALGGQDGSCHPAAPGLVGPAPRTPGVGRIWVRSSIPSH 127
 QY 127 AASPTSGTYRSSAPLLEALPGPWRMASGFWKTA 160
 Db 128 AASPTSGTYRSSAPPGWALPGPWHMVSGFWRKA 161

RESULT 12
 US-10-601-020-13
 ; Sequence 13, Application US/10601020
 ; Publication No. US20040156862A1

Query Match 63.8%; Score 547; DB 9; Length 143;
Best Local Similarity 78.5%; Pred. No. 3.5e-42;
Matches 106; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

Qy	7	KKKTNVPTVAHRTSSSRVAVRSLVEFTCCRAGALDWVCARRRLPGRNLEVDVLSLSPR	66
Db	8	QRKPNVTPTAAHRTSSSRVAVRSLVEFTCCRAGAPGWVCARLGRLPGRNLEVDNLSPLR	67
Qy	67	LVGFRAGPGLSPGTLGFSMAMRAAGGDSCLPVALGLAGAPOTPGVGRRAIWVRSIPLR	126
Db	68	LASFRAGPGLSPGTPGFSMAMRAWGGDSCHPAAPGLVGAPKTPGVGRVIVVRSIPLSH	127
Qy	127	AASPTSWGTYRSSAP	141
Db	128	AASPTSWGTFRSSAP	142

Search completed: August 25, 2005, 03:29:13
Job time : 162 secs